TR-I-0167

AUTOMATIC LANGUAGE RECOGNITION USING ACOUSTIC FEATURES

杉山 雅英

Masahide SUGIYAMA

概要

Language recognition (e.g. Japanese, English, German, etc) using acoustic features is an important yet difficult problem for current speech technology. In this report, two language recognition algorithms are proposed and some experimental results are described. The speech data base used in this report contains 20 languages. The speech data was carefully divided into training and test sets, recognition experiments being designed as both speaker-independent and text-independent. The first algorithm is based on the standard Vector Quantization (VQ) technique. The second algorithm is based on a single universal (common) VQ codebook for all languages, and its occurrence probability histograms. The experimental results show that the recognition rates for the first and second algorithms were 65% and 80%, respectively, each using just 8 sentences of unknown speech (about 64 seconds). With sufficient input speech the second algorithm is better than the first.

ATR Interpreting Telephony Research Labs.

ATR 自動翻訳電話研究所

Contents

1 Introduction

1.1 Why this research

Recently a lots of research have been done on speech recognition[1]. Automatic language recognition can be very useful for other speech recognition systems. This could allow speech recognition systems to be used in a more general environment. But most of this systems require identification to be perform on a few seconds and noisy environment.

1.2 Previous research on the same subject

A few research has been conducted on this subject such as the work of Foil[2]. He developed two algorithms. The first one was based on Pitch and Energy features and the second one was based on formant Clustering. This tests has been performed on 3 languages with approximately 6 hours speech for training and 4 hour of non proceed speech for final tests. The averaged test duration was about 4.5 sec and the signal to noise ratio (SNR), 5dB. The conclusion was that formant clustering method performed well (in Table 1).

Following this study a second one has been conducted by Goodman, Martin and Wahlfor [3] to enhance the previous formant clustering algorithm. The algorithm based on LPC based formant extraction, was greatly improved, more than 50%. Test duration were less than 10 sec. with a 9 hours data base, an average Signal to Noise Ratio of 9 dB and 6 languages (in Table 2).

Table 1: J.T.Foil's study

language	3 languages (pick up in major language family)
speech data	training 2 hours per language, test 4 hours
methods	1. prosodic feature (pitch, energy) & quadratic classifier
	2. VQ distortion for formant information in voiced-part
performance	64% identification rate(11% reject)
note	1. noise environment (averaged SNR 5dB)
	2. short test input utterance (4.5sec)

Table 2: F.Goodman's study

language	6 languages
speech data	more than 9 hours
methods	improved Foil's formant based algorithm
performance	improvement: more than 50%
note	1. noise environment (averaged SNR 9dB)
	2. short test input utterance (less than 10sec)

1.3 Introduction to our work

This study used completely different acoustic feature — The first idea was to try methods used on speaker identification[4],[5] and speech recognition for language recognition.

1.4 Structure of the report

The Chapter 2 of this report describes the speech data base used and the choice for training and tests. Comparatively to previous studies our data base is very small and with low noise. In Chapter 3, two algorithms based on

Vector Quantization have been tried, using several distances (CEP, WLR, CEP, CEP+ΔCEP, WCEP, FWCEP) between speech wave acoustic features. The Chapter 4 presents the results obtained with each method, their analysis and further efforts that could be tried.

2 Data

2.1 Description

The database is a 20 languages database[18]. The table in Appendix A describes for each language the administration who produced the files, the number of speakers and the number of sentences. The original sampling frequency is 32 kHz but all files have been converted to 8 kHz.

An important distinction must be made between sentences and files. In this data base, a file can have several sentences and the number of sentences per file is not constant. (see Appendix I) A file is the minimum set of data that can be separated for training or test.

For our experiments, only AT&T files have been used for AM and for 2 languages (FI and HI) all files have not been used. In all further description, these eliminated files do not appear. See Appendix H

The Table 10 in Appendix E describes the average and total duration of speech files. In Appendix J.2, the sound spectrograms for several languages have been displayed.

2.2 Choice of data

The precise choice of files from the data base is described in Appendix H. The data base, after few elimination of files has been split into two separate parts:

- training data
- testing data

The amount of data for each part is approximately the same: 20.9 min for training and 20.5 min for test. Table 10 shows the amount of data and the average duration for training and test.

All languages have different speakers and different sentences for training and test, excepted:

- AM which has 4 sentences in each file.
  - same sentence for training and tests but with different speakers.
- DA, FR and JA which have only two sentences.
  - two same sentences are used for both test and training but different speakers for test and training.

All languages have the same number of male and female speakers for both test and training excepted HI which has:

training	3 males & 4 females
test	3 males & 3 females

2.3 Problems summary

This data base is from many viewpoints not good for language recognition experiments as shown in Table 3:

1. Number of sentences in one file

   As described before, some files contain several sentences: this is the case for: AM, DA, EN, FR, HU & IT.

2. Number of files

   As described in Table 10 all languages have 16 files, 8 for training and 8 for test excepted AM (4 for training and 4 for test) and HI (7 for training and 6 for test).

3. Number of different sentences

   AM, DA, FR and JA have only 1 or 2 sentences and for this reason same sentences are used for training and test.

4. Duration of each file

   Table 10 shows that AR, CH, EN, GE, HI, have shorter files than the average and AM, DA, FR, longer.

5. amount of data

   Table 10 shows that AM, AR, CH, AN, GE, and HI have a particularly small amount of data and DA & FR a particularly important one.

6. Signal to noise ratio(SNR)

   The data base has a very high signal to noise ratio (very low noise).

   It can be

   - a good thing to test and optimize algorithm with good quality speech.
   - a bad thing because the algorithm can be inefficient in noisy environment which is actually the real environment in which this kind of system is to be used.

   In Table 11, GE, PR & SP has been classified has being noisy comparatively to other languages but in fact they are very clear comparatively to other experiment conducted on language recognition[2],[3].

Table 3: Problems summary

language	number of sentences in each file	number of files	number of sentences	duration S & L	amount of data	SNR
AM	x	x	x	L	x	
AR				S	x	
CH				S	x	
DA	x		x	L		
DU						
EN	x			S	x	
FI						
FR	x		x	L		
GE				S	x	x
GR						
HI		x		S	x	
HU	x					
IT	x					
JA			x			
NO						
PL						
PR						x
RU						
SP						x
SW						

S (short); L (long)

3 Recognition Methods

3.1 Vector Quantization (VQ)

Since VQ is a clustering algorithm the general idea that in a k dimensional space with a given set of vectors we can find clusters of vectors. Each cluster is replaced by its 'centroid', which means the most representative vector of the cluster. So q vectors can be represented, quantized, by p centroids ($p < q$) of clusters. These centroids are called codes and the set of p codes is called the codebook.

VQ algorithm is based on the principal of minimization of distortion. Let Ω be a set of training vectors and V be a vector codebook which is a set of representative vectors of Ω. The method used to generate V is base on the VQ distortion $D(\Omega, V)$ minimization defined as follows:

$$D(\Omega, V) = \sum_{x \in \Omega} d^2(x, V) \tag{1}$$

$$d(x, V) = \inf_{v \in V} d(x, v) \tag{2}$$

Eq.(2) means the element x is quantized (or replaced) by v, which is the nearest to x. Eq.(1) means that the distortions for all elements in Ω are accumulated. Usually, the number of elements in V is about 10% of that in Ω. The algorithm for generating V from Ω has been formulated by Linde, Buzo and Gray. Their algorithm is not the optimal (suboptimal). It can provide the sequence of V_n and $D(\Omega, V_n)$ is a decreasing sequence. In Appendix J.3 the generated vecor codes are displayed.

3.2 Acoustic parameters

A wide variety of acoustic parameters can be defined from a given speech wave frames[16]. We used the following feature parameters:

autocorrelation coefficient
LPC cepstrum coefficient
Δ cep coefficient

Table 4: Specification for speech analysis

autocorrelation analysis order	13
LPC analysis order	10
frame length	128 (16ms)
sampling frequency	8 kHz
VQ relative distortion	0.01

Fig.1 shows the speech analysis procedure. Here the auto-correlation method is used to compute LPC coefficient.

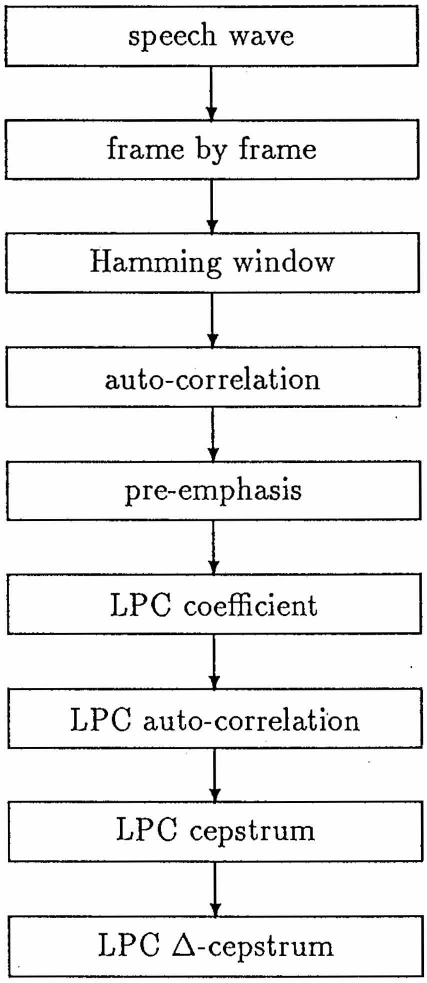


Figure 1: The Diagram for Speech Analysis Procedure

3.2.1 Speech wave and frames: (x_i)

The speech wave is a time series of n value x_i, n is the length of the file containing the speech wave. In our experiments, the frequency of time series is 8 kHz. A frame is a subset of a speech wave:

$$f_i^j = x_{js+i} \qquad i = 1, \dots, l \tag{3}$$

define the l-th sample of j-th frame, where s is the shifting between two frames, l the length of the frame.

3.2.2 Hamming window

To eliminate hard discontinuities at the border of the frame, which can induce problems in frequency domain, we are using a Hamming window to smooth them.

3.2.3 Auto-correlation coefficient: (r_i)

Normalized auto-correlation coefficients are computed as follows;

$$\rho_k = \sum_{n=1}^{l-k} x_n x_{n+k} \qquad k = 0, \dots, m \tag{4}$$

$$s_k = \rho_k / \rho_0 \qquad k = 1, \dots, m \tag{5}$$

l: length of the frame, k: auto-correlation order
Here we used $m = 13$.

3.2.4 Pre-emphasis

Pre-emphasis is a kind of spectral flattening done using a filter which transfer function is:

$$H(z) = 1 - 0.98z^{-1} \tag{6}$$

The pre-emphasis was performed in the auto-correlation domain.

3.2.5 LPC coefficient: (a_i)

LPC analysis is a modeling of speech wave form by Auto Regressive (AR) model. It means that speech wave $x_i\{i = 1, \ldots, n\}$ is modeled at time t by:

$$\hat{x}_t = -a_1 x_{t-1} - a_2 x_{t-2} - \ldots - a_p x_{t-p} \tag{7}$$

p is the order of LPC analysis, we used $p = 10$. This is called Linear Prediction because time t is defined from time $t-1$ to $t-p$. a_i are called LPC coefficient and calculated from the auto-correlation coefficient.

3.2.6 LPC auto-correlation coefficient: (r_i)

These coefficient are extended correlation coefficient using LPC coefficients.

$$r_i = \sum_{k=1}^{p} a_k r_{i-k} \qquad (i > p) \tag{8}$$

Here, we used $n = 16$.

3.2.7 LPC cepstrum coefficient: (c_i)

$$\begin{aligned} c_1 &= -a_1 \\ c_i &= -a_i - \sum_{k=1}^{n-1}(1 - \frac{k}{n}) a_k c_{n-k} \qquad (2 \leq n \leq p) \\ c_i &= -\sum_{k=1}^{n-1}(1 - \frac{k}{n}) a_k c_{n-k} \qquad (p < n) \end{aligned} \tag{9}$$

These coefficient contain instantaneous spectral information which is used by human being in speech perception.

3.2.8 Δcep coefficient: (Δc_i)

Δcep coefficients are a kind of derivative of c_i coefficient and computed as the slope of the approximate line for

$$c_i(t-q), \ldots, c_i(t), \ldots, c_i(t+q) \tag{10}$$

$c_i(t)$ is the i-th cepstral coefficient at time t. These coefficients represent dynamic spectral features (spectral transition) which are believed to play important role in human speech perception.

3.3 Several distances

3.3.1 Cepstral distance (CEP)

The cepstral distance[6] is the mean square integration between two logarithmic spectra:

$$d^2_{CEP}(f,g) = \int_{\pi}^{\pi} (\log f(\lambda) - \log g(\lambda))^2 \frac{d\lambda}{2\pi} \tag{11}$$

In our experiments the following equation is used:

$$d^2_{CEP}(f,g) = \sum_{i=1}^{N} (c_i^{(f)} - c_i^{(g)})^2 \tag{12}$$

where $(c_1^{(f)}, \ldots, c_n^{(f)})$ and $(c_1^{(g)}, \ldots, c_n^{(g)})$ are two vectors of cepstral coefficients. We used $N = 16$.

3.3.2 Weighted Likelihood Ratio (WLR)

Since an important information of speech is located in spectral peaks, this distance was developed to be more sensitive to these peaks (formants): recognition[10],[13].

WLR measure is defined as follows:

$$d^2_{WLR}(f,g) = \int_{\pi}^{\pi} (f(\lambda) - g(\lambda))(\log f(\lambda) - \log g(\lambda)) \frac{d\lambda}{2\pi} \tag{13}$$

In our experiments, we used the truncated formula:

$$d^2_{WLR}(f,g) = \sum_{i=1}^{N} (r_i^{(f)} - r_i^{(g)})(c_i^{(f)} - c_i^{(g)}) \tag{14}$$

Here, we used $N = 16$.

3.3.3 ΔCEP distance (Δ CEP)

ΔCEP distance can be defined as Euclidean distance between Δcepstrum coefficients. We used the truncated form;

$$d^2_{\Delta CEP}(f,g) = \sum_{i=1}^{N}(\Delta c_i^{(f)} - \Delta c_i^{(g)})^2 \tag{15}$$

where, $N = 10$. ΔCEP characterizes the difference between two transitional spectra.

3.3.4 CEP + ΔCEP distance

Since these two distances should contain different and complementary information. We defined CEP + ΔCEP distance as;

$$t\frac{1}{\sigma_{CEP}}\sum_j d^2_{CEP} + (1-t)\frac{1}{\sigma_{\Delta CEP}}\sum_j d^2_{\Delta CEP} \tag{16}$$

where t is a weighting factor and σ_{CEP} and $\sigma_{\Delta CEP}$ are averaged values of CEP and ΔCEP.

3.3.5 Weighted Cepstrum distance (WCEP)

This distance has been proved to be efficient in some speech recognition application[15].

$$d^2_{WCEP}(f,g) = \sum_{i=1}^{N} w_i(c_i^{(f)} - c_i^{(g)})^2 \tag{17}$$

$$w_i = \begin{cases} i^2 & (1 \le i \le i_0) \\ i_0^2 & (i_0 < i) \end{cases} \tag{18}$$

Here, $N = 10, i_0 = 6$.

3.3.6 Frequency Weighted Cepstrum distance (FWCEP)

As the human being are more sensitive to low frequencies [12],[13], FWCEP is defined as:

$$d^2_{FWCEP}(f,g) = \int_{\pi}^{\pi} (\log f(\lambda) - \log g(\lambda))^2 w_F(\lambda) \frac{d\lambda}{2\pi} \tag{19}$$

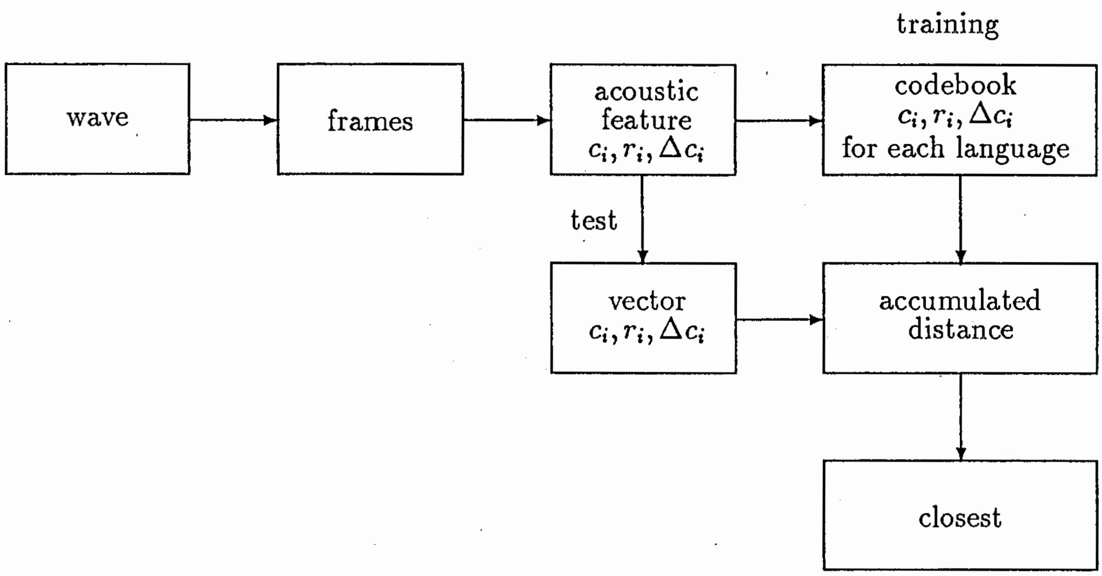


Figure 2: Concept of standard VQ algorithm

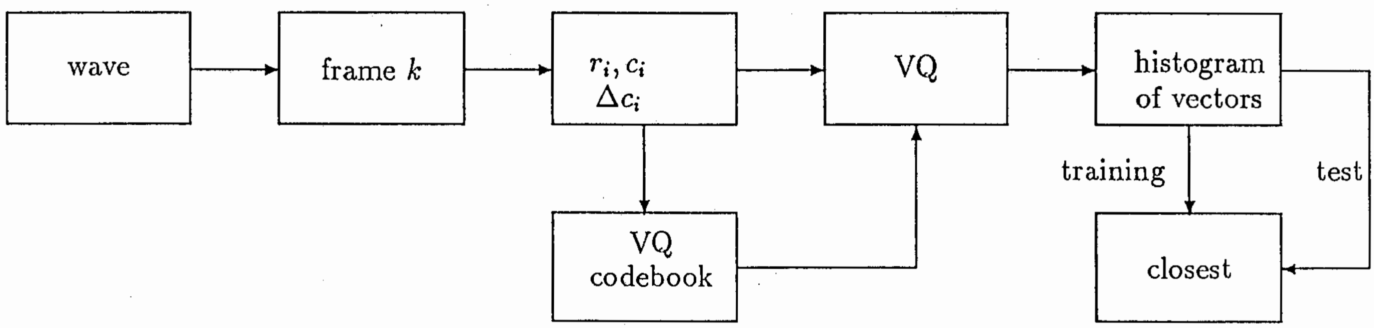


Figure 3: Concept of Histogram VQ algorithm

where $w_F(\lambda)$ is a frequency weighting function[11].

3.4 Recognition algorithm

3.4.1 Standard VQ algorithm

In this case, a code book is generated for each language using all training sentences. According to the distance which is chosen, correlation, LPC cepstrum or Δcep coefficients are computed and a code book is made for every needed coefficients. The accumulated distance is defined as;

$$d_k = \sum_{i=1}^{n} \min_j d^2(v_i, V_j^k) \tag{20}$$

where v_i is a testing vector and V_j^k is the j-th vector of k-th codebook. The distance d can be any distance explained before but of course it must be the same as the one used for codebook generation. Each language is characterized by its VQ codebook.

3.4.2 VQ histogram algorithm

In this case, an universal code book is generated for all training data. Universal code book:

$$\{u_1, \ldots, u_q\} \tag{21}$$

After making this universal code book, each language is characterized by the occurrence probability histogram $h_1, \ldots, h_q$ of each vector u_i of the universal code book for this language. h_i is the occurrence probability of vector u_i. Each testing sentence is also characterized by its occurrence probability histogram. The distance between a test sentence and a reference language is the Euclidean distance between the histograms. This similar idea has been already applied to speaker identification problem[20].

4 Experimental Results

4.1 Result computation

For a given input test file (or a set of files) F, the following items are computed:

4.1.1 Distance to each language L: $D(F, L)$

For each input test file, a distance is computed between this file and each reference language:

$D(F, L)$ is the accumulated distance between frames of F and a codebook of L in the case of standard VQ algorithm. $D(S, L)$ is the Euclidean distance between histogram of F and histogram of L in the case of VQ histogram algorithm.

4.1.2 Recognition rank: $R(L, F)$

According to these distances, all languages are sorted from minimum distance to maximum. $R(L, F)$ is the rank of language L after sorting.

4.1.3 Percentage of success

For each language, 20 percentage of success are computed as follows;

$$choice_n(L) = \frac{\sum_{i=1}^{T} N[R(L, F_i) = n]}{T} \tag{22}$$

where F_i is input test file i for language L, T is the total number of tests for language L and N is the function defined as $N(Q) = 1$ if Q is true, 0 if Q is false

4.1.4 Total percentage of success

For each choice a total percentage of success is computed;

$$choice\_n = \frac{\sum_j \sum_{i=1}^{S} N[R(L_j, F_i) = n]}{S} \tag{23}$$

where $S = \sum_j T_j$

4.1.5 Confusion matrix

The confusion matrix is the average distortion between the frames of language n and reference of language m. Bigger rectangles show low distortion (close to) and small rectangles show high distortion (far from). Horizontally is shown the average distortion of an input sentence and the 20 code books.

4.2 Standard VQ algorithm results

4.2.1 Size of the codebook

The influence of the size of codebooks has been studied with WLR distance, a shifting between training frames of 128 samples, a shifting between test frames of 512 samples. The average number of training frame was 3900 (minimum for AR: 2260, maximum for FR: 6104).

Fig.4 shows that from 16 codes (included) the result are nearly stable and under 16 the results are going down. One explanation might be that the amount of training data is not big and may be not so representative of all feature of each language. But the number of language is 20 and the number of important generated vector is between 8 and 16. So it might mean that each language does not have a specific important acoustic cluster but probably share them with other languages. If we look up to the third choice results (top third recognition rate) are stable from 32 vector codes.

We chose to use 16 codes for the following experiments but distances comparison, since the size of the code books has been studied with only one distance, has been tried with also 64 codes in order to be sure we are in the stable part.

4.2.2 Shifting between training vectors

The influence of the shifting between training vectors has been studied with WLR distance, a shifting between testing frame of 512 samples and 16 vector codebook. Fig.5 shows that the accuracy is the best for a shifting of 128. It seems normal that over 128, as the number of frame is lower, the accuracy decreases. But under 128 as the number of frames increases the accuracy is going down. One explanation is probably that as our frame length is 128; under 128 sample shifting we use redundant information.

4.2.3 Shifting between testing vectors

As for the two previous case this has been studied with WLR distance. A shifting of 128 was used between frames to generate codebooks and the code book size was 16. From Fig.6 we can see the accuracy is constantly increasing with the number of frames (when the shifting is lower). Under 64, the accuracy seems to stabilize so 64 was chosen for next experiments.

4.2.4 CEP + ΔCEP parameter

In the case of CEP + ΔCEP, cepstrum and Δcepstrum codebook are generated separately. During tests the accumulated distance used is Eq.(16). Fig.7 shows that CEP + ΔCEP is better than CEP alone or ΔCEP alone. This means that instantaneous frequency parameter and transitional frequency parameter are complementary. However, frequency variation parameter used alone (ΔCEP) has very poor result. It would mean more information is contained in instant frequency than in frequency variation for language identification by this algorithm. The best value for t proved to be 0.5 and will be used in next experiments.

4.2.5 Different distances comparison

Fig.8 represents the percentage of success for 6 distances; CEP, WLR, ΔCEP, CEP + ΔCEP, FWCEP and WCEP in function of the amount of input data for each test. The number indicated for the amount of data are corresponding to the number of files taken together for 1 test.

Here the shifting used for codebook was 128 (average of 3800 frames) and the shifting used for test was 64. The size of the codebook was 16. The same test was conducted on four distances (CEP, WLR, ΔCEP & FWCEP) with 64 codes and the results proved to be 5 to 10% better. This Fig.9 shows that the best distance is always WLR with the best results when the amount of data is more important. CEP has a similar behavior but with lower results. FWCEP is the same as CEP in amount of data 1 case but better with bigger amount of test data.

ΔCEP did not perform well and this confirms previous test on CEP + ΔCEP. This would mean that the acoustic information is more in frequencies than in their transitions. And WLR's good results seems to indicate that the most important information is contained in the peaks in instantaneous frequency domain. FWCEP results seems to show that frequency weighting can improve accuracy and frequency weighted WLR might perform well, but has not been tried.

If we compare recognition accuracy from Fig.10 for different distances, it appear that with different distances, errors are made on same languages (comparison made with four best distances; WLR, FWCEP, CEP + ΔCEP , CEP and the amount of data is 4). It would mean that different distance does not recognize better same languages than the others. It is shown also by the confusion matrix which are very similar for two different distances.

4.2.6 Confusion matrix analysis

The confusion matrix Fig.11 & Fig.12 show that some languages such as AM, DA or PL highly confuse with other languages. This could mean that some languages have no specific acoustic features in the sense of our experiments. However these three languages performed very well.

If we try to compare the linguistic classification of languages (AppendixG) and the confusion matrix it seems that our acoustic classification is different. For example, Uralic & Altaic languages (FI, HU and JA) are not specially misrecognized as Uralic & Altaic.

The case of AR is interesting. It has its own language family (Semitic & Hamitic) but has poor recognition accuracy. Its codebook is far from most of other languages tests but its test are not close to its codebook. It might come from the asymmetricity of the accumulated distance. But if we look at PR which codebook is also far from most of the other language tests, its test are close to PR.

4.3 VQ histogram algorithm results

4.3.1 Size of the universal codebook

As for standard VQ algorithm this experiment was done WLR. The shifting for training was 1092. This implies about 10000 frames for universal code book. The shifting for test was 512 and 1 file was taken for each test. The shape of result (Fig.13) is close to one obtained with standard VQ algorithm. In this case a peak appeared for 32 codes with the first choice but this peak is located at 64 codes for the top second and top third case. In further experiment we will use 32 codes but the 64 case will be tried for distance comparison.

4.3.2 Shifting for codebook generation

The results (Fig.14) are similar to the results obtained with standard VQ algorithm (Fig.5). The best result is for a shifting of 128 and redundant information gives worse results. Here we WLR, 32 vectors for the universal code book and a shifting of 512 samples between testing frames.

4.3.3 Shifting for tests

The result (Fig.15) is also close to the result (Fig.6) obtained with the standard VQ algorithm, and in the case of shape, small shifting, under the length of the frame (means redundant) can be profitable. The parameter used here were amount of data is 1, frame shifting for training 128 and the distance was WLR.

4.3.4 Different distances comparison

Training and test shifting used were the same as standard VQ algorithm: 128 for training and 64 for test.

2 series of results have been computed:

- 32 vectors universal code book Fig.16
- 64 vectors universal code book Fig.17

The size of the bigger code book did not have better results and best result was even obtained with 32 vectors universal code book. The most important thing is that according to Fig.16 & Fig.17 this algorithm have not so different results with different distances, excepted maybe WCEP which did not performed well. One explanation might be that there are 2 different concepts used for distance:

- the distance between acoustic parameters
- the distance between histograms.

We can also notice that best results are obtained when same distance is use (Euclidean in our case) between acoustic parameters and histogram.

4.3.5 Robustness in noisy environment

We conducted this test assuming that in noisy environments, mainly low energy frames are affected. Thus, we computed for each files the maximum power M and eliminated all frames having a power under $M - x$ dB. We tried $x = 50$ dB and $x = 40$ dB just in one case with VQ histogram algorithm. The table 5 shows that our algorithm is probably not very robust to noisy environment.

number of codes	64
distance	WLR
amount of test data	8
frame shift for training	1092
frame shift for test	512

Table 5: Robustness in noisy environment

SNR	recognition accuracy
$-\infty$	75 %
50 dB	65 %
40 dB	45 %

4.4 Comparison between two methods

As shown the result Fig.8 and Fig.9 and result Fig.16 and Fig.17, the results obtained with VQ histogram are better than with the standard VQ algorithm. In both case, the influence of training and test shifting between frames is the same: for training best results are obtained when all information is used but with no redundancy and for tests results always improved when the shifting decrease. For the standard VQ algorithm and VQ histogram algorithm the shifting for the test is 64 and for training 128. The number of codes are 16 (Fig.8) and 64 per language (Fig.9) in the case of standard VQ algorithm; 32 (Fig.16) and 64 for universal codebook (Fig.17) in the case of VQ histograms algorithm.

It is quite surprising to see that the best distance for result (Fig.16) is CEP and the it has poor result with standard VQ algorithm. FWCEP is second but as in Fig.8 and Fig.17 it is under CEP. WLR performed well in both case. ΔCEP has better results, relatively to other distance with VQ histogram algorithm.

This might be explained by the fact the universal codebook is made out of 20 time more data than each codebook for standard VQ algorithm and then might have more representative codes. Another reason could be that the distance defined between algorithms is more adequate to this kind of problems than the accumulated distance or that histogram contain better information for language recognition.

Anyway, we can say according to the results that first algorithm has big difference between results obtained with different distances whereas the second one has not so big different results with different distances.

5 Other Methods and Further Efforts

5.1 Neural Network based algorithm

We began to try another algorithm base on neural network. In this case the acoustic feature used to characterize each reference language or sentence was a probability histogram of the value of the maximum auto-correlation or modified auto-correlation. This means that for each frame maximum auto-correlation or maximum modified auto correlation was extracted from the pitch interval and the probability histogram was generated.

This histograms were fed in a neural network and the output were the distortion to each language. This algorithm did not performed well and the two main reasons may be:

- we did not have enough data to use neural network.
- these histogram did not contain adequate information for languages recognition.

However going further on in this direction might be interesting. We plan to try to use pitch and energy features in further research.

5.2 Identification of written languages

From the transcription of the sentences we tried a very simple algorithm to try to identify written languages. For each reference language or test sentence we made a symbol occurrence probability histogram, where symbol means letters from 'a' to 'z' and spaces. ',', ';', ':' were considered has 2 spaces, other symbols ere ignored. For 13 languages we obtained 91 % accuracy with 2 sentences for each test and 100 % with 8 sentences for each test.

5.3 Future studies

We have the following items for future studies:

- Histogram smoothing technique
- Distance measure for two histograms
- pitch histogram and Euclidian distance method

6 Conclusions

According to the previous analysis we can conclude that:

- for both algorithm tried:
  - The accuracy does not necessarily increase when the number of codes for codebook does: there seems to be an optimal codebook size.
  - The shifting between frames for training gives the best results when it is equal to the frame length: means when all information available is used with no redundancy.
  - The accuracy increase with the number of frames used for test: in this case redundant information can be beneficial.

- The VQ histogram algorithm proved to perform better than the standard and with different behavior in function of the distances used:
  - the standard VQ algorithm have different results with different distances and best results are obtained with WLR.
  - the VQ histogram algorithm have similar results with different distances.Best result is with CEP.
- The acoustic classification seems to be different of the linguistic classification.
- Our algorithm seems to be not very robust in noisy environment.

7 Acknowledgments

This study was done by Paul Hiriart(INSA, Lyon) under the author's supervising at NTT Basic Research Labs. The author acknowledge the suggestions, contributions made by Dr. Honda, Head of Information Science Dep., for his useful comments and his down sampling tool, Mr. Shiraki for his laser printer output tool, Dr. Kawahara, Mr. Matsuoka, Mr. Uchiyama for their suggestions on neural networks, Mr. Itoh for his pitch extraction algorithm, Dr. Mano and Miss Matsui for many suggestion. Also I would like to thank Dr. Osumi for his help, Mr. Kakehi, Executive Manager of Information Science Dep., and Dr. Katoh, Exectutive manager of NTT basic Research Labs., for their guidance and encouragement. Finally I would like to thank all the person who do not appear hereby and helped me in many ways.

References

[1] S.Furui, "Digital Speech Processing, Synthesis, and Recognition", Marcel Dekker Inc.

[2] J.T.Foil, "Language identification Using Noisy Speech", ICASSP86, 17.1 (1986-03).

[3] F.Goodman, A.Martin, "Improved Language Identification in Noisy Speech", ICASSP89, 35.S10b.4 (1989-05).

[4] F.K.Soong, A.E.Rosenberg, "On the Use of Instantaneous and Transitional spectral Information in Speaker Recognition", ICASSP86, 17.5 (1986-04).

[5] M.Sugiyama, "Segment based Text Independent Speaker Recognition", ASJ pp.75 (1988-03).

[6] B.S. Atal, "Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker Identification and Verification", J. Acoust. Soc. Am., 55, pp.1304-1312 (1974).

[7] F. Itakura, "Minimum Prediction Residual Principle Applied to Speech Recognition", Trans. IEEE, ASSP-23, pp.67-72 (Feb. 1975).

[8] A.H. Gray and J.D. Markel, "Distance Measures for Speech Processing", IEEE Trans., ASSP-24, 5, pp.380-391 (Oct. 1976).

[9] R.M. Gray et al, "Distortion Measures for Speech Processing", IEEE Trans., ASSP-28, 4, pp.367-376 (Aug. 1980).

[10] M. Sugiyama and K. Shikano, "LPC Peak Weighted Spectral Matching Measures", Trans. of IECE, Vol.J64-A, No.5, pp.409-416 (May 1981) (in Japanese).

[11] M.Sugiyama and K.Shikano, "Frequency Weighting in LPC Peak Weighted Spectral Matching Measures", Proc. of IECE, 1352 (Apr. 1981) (in Japanese).

[12] M. Sugiyama and K. Shikano, "Frequency Weighted LPC Spectral Matching Measures", Trans. IECE, Vol.J65-A, No.9, pp.965-972 (Sep. 1982) (in Japanese).

[13] M. Sugiyama and K. Shikano, "WLR Measure Applied to Word Recognition", Trans. of IECE, Vol. J66-D, No.4, pp.385-391 (Apr. 1983) (in Japanese).

[14] M. Sugiyama and K. Shikano, "Optimization of Peak Weighting in Weighted Likelihood Ratio", Proc. of ASJ, 1-1-12, pp.381-382 (Oct. 1980) (in Japanese).

[15] Y. Tohkura, "A Weighted Cepstral Distance Measure for Speech Recognition", ICASSP86, 14.17, pp.761-764 (Apr. 1986).

[16] J. D. Markel and A. H. Gray, "Linear Prediction of Speech", Springer, 1976.

[17] S.Nishimura, etal, "Speaker Recognition with Neural Networks (1) & (2)", ASJ pp.53-56 (1989-03).

[18] Irii, Itoh, etal, "Multi-lingual Speech data base for Speech Quality Measurements and its Statistic Characteristics", SP87-69 (1987).

[19] K.Itoh, N.Kitawaki and K.Kakehi, "A Study of Objective Quality Measures for Digital Speech Waveform Coding Systems", Trans. of IECE, Vol.J66-A, No.3, pp.274-281 (Mar. 1983) (in Japanese).

[20] K.Shirai, K.Mano and S.Ishige, "Speaker Identification based on Frequency Distribution of Vector-Quantized Spectra", Trans. of IECE, Vol.J70-D, No.6, pp.1181-1188 (Jun. 1987) (in Japanese).

List of Tables

List of Figures

A The Number of Speakers and Sentences

Languages	Administration	Speakers	Sentences
American English (AM)	AT&T	Male 4	1(4)
		Female 4	1(4)
	BNR	Male 2	4(4)
		Female 2	4(4)
Arabic (AR)	CCITT Lab.	Male 2	4
		Female 2	4
Chinese (CH)	Institute of Telecom Transmission in Beijing	Male 4	8
		Female 4	8
Danish (DA)	KTAS	Male 4	2
		Female 4	
Dutch (DU)	DNL	Male 4	8
		Female 4	8
English (EN)	BT	Male 8	8(3)
		Female 8	8(3)
Finnish (FI)	Posts and Telecom. of Finland	Male 4	8
		Female 4	8
French (FR)	CNET	Male 4	2
		Female 4	2
German (GE)	FTZ	Male 4	8
		Female 4	8
Greek (GR)	CCITT Lab.	Male 2	4
		Female 2	4
Hindi (HI)	CCITT Lab.	Male 3	3
		Famale 2	4/3
Hungarian (HU)	CCITT Lab.	Male 4	8
		Famele 4	8
Italian (HU)	CSELT	Male 4	8(4)
		Famele 4	8(4)
Japanese (JA)	NTT	Male 4	2
		Famele 4	2
Norwegian (NO)	Norwegian Telecom. Administration	Male 6	2
		Famele 6	2
Polish (PL)	CCITT Lab.	Male 4	8
		Famele 4	8
Portuguese (PR)	EMBRATEL	Male 4	8
		Famele 4	8
Russian (RU)	Ministry of Posts and Telecommunication	Male 4	8
		Famele 4	8
Spanish (SP)	Telephonica	Male 4	8
		Famele 4	8
Swedish (SW)	Swedish Telecom. Admin	Male 4	8
		Famele 4	8

B LPC distance and Calculation Formula for VQ based algorithm

Table 6: LPC distance for VQ based algorithm

distance	name
CEP	Cepstral distance
WLR	Weighted Likelihood Ratio
Δ CEP	Δ cepstrum distance
Δ CEP + CEP	combination of Δ cep and cepstrum distance
FWCEP	Frequency Weighted Cepstral distance
WCEP	Weighted Cepstral distance

Table 7: Formula for several distance calculation

distance	formula
WLR	$\sum_{n=1}^{N}(r_n - r_n^*)(c_n - c_n^*)$
CEP	$\sum_{n=1}^{N}(c_n - c_n^*)^2$
FWCEP	Frequency Weighted Cepstral distance
WCEP	$\sum_{n=1}^{N} w_n(c_n - c_n^*)^2$

N; truncation order (N=16)

C Linguistic Classification of 20 Languages

Table 8: Linguistic classification of 20 Languages

Semitic & Hamitic		Arabic
Sino-Tibetan		Chinese
Uralic & Altaic		Finnish
		Hungarian
		Japanese
Indo-European	Greek	Greek
	Indo-Iranian	Hindi
	Germanic	(west)
		American English
		Dutch
		German
		(north)
		Danish
		Norwegian
		Swedish
	Romance	French
		Italian
		Portuguese
		Spanish
	Slavic	Polish
		Russian

D Abbreviation of Each Language

Table 9: Abbreviation of each language

LANGUAGE	abbreviation
American	AM
Arabic	AR
Chinese	CH
Danish	DA
Dutch	DU
English	EN
Finnish	FI
French	FR
German	GE
Greek	GR
Hindi	HI
Hungarian	HU
Italian	IT
Japanese	JA
Norwegian	NO
Polish	PL
Portuguese	PR
Russian	RU
Spanish	SP
Swedish	SW

E Duration Time of Each Language

Table 10: Duration time of each language

LANGUAGE	number of files	average duration (file) ms	average duration (language) ms
AM	8	10.5**	84.1 †
AR	16	4.8*	77.5 †
CH	16	6.1*	98.0 †
DA	16	11.6**	177.2 ‡
DU	16	8.3	132.6
EN	16	6.2*	100.0 †
FI	16	8.4	135.1
FR	16	11.8**	204.1 ‡
GE	16	6.0*	95.4 †
GR	16	9.4	150.6
HI	13	7.2*	93.1 †
HU	16	7.8	123.9
IT	16	8.7	139.3
JA	16	9.5	152.0
NO	16	7.3	115.7
PL	16	8.6	136.6
PR	16	7.2	116.1
RU	16	7.6	121.2
SP	16	7.5	120.1
SW	16	7.1	111.0
average		8.1	124.2

* low, ** high; † low, ‡ high

F Signal to Noise Ratio of Each Language

Table 11: Signal to Noise Ratio of each language

LANGUAGE	SNR
AM	49 (40)
AR	51 (49)
CH	49 (47)
DA	48 (47)
DU	56 (55)
EN	51 (48)
FI	57 (55)
FR	60 (58)
GE	36 (34)
GR	51 (48)
HI	48 (46)
HU	45 (41)
IT	56 (55)
JA	61 (57)
NO	44 (43)
PL	47 (46)
PR	38 (33)
RU	49 (42)
SP	39 (32)
SW	56 (46)

SNR: A weighting is used.
the value in ( ) is for no weighting case

G Linguistic Classification by Human Listening

In order to determine the classification of languages, the human listening test is carried out. The listeners are 4 female Japaneses and 1 male French. Japanese listers are trained listening crews. Pairs of two sentences are presented to listeners, and each listener respond a value which corresponds the nearness between two sentences in the sense of language. When the listener is presented a pair of sentences, they must judge whether the two sentences belong to the same language or not. When they judge as the same language, the value is 0. Otherwise, the value is 1. When the listener knows one of two languages in the pairs, the value may be 0.

The number of the sentence pairs is 380 (= 19x19 + 19). The listeners are very familiar with Japanese and it is trivial for them to distinguish Japanese from the other languages, then language Japanese is excluded in the listening test. Therefore, the number of languages is 19 and the number of all combination is 19x19. In order to increase the number of diagonal trial, which two sentences are in the same language, these diagonal 19 pairs are duplicated. These all pairs are randamized. A series of sentence pairs are divided into each 10 pairs, and a long beep signal is placed.

The following table is the confusion matrix by the human listening test. The vertical line and the horizontal line in the table correspond the the first sentence and the second sentence in each presented pair. The each value is normalized the total presented number. When the i-th row, j-th column element is close to 0, this means i-th and j-th languages are close in the sense of acoustic features.

Table 12: Confusion matrix by human listening test

	am	ar	ch	da	du	en	fi	fr	ge	gr	hi	hu	it	no	pl	pr	ru	sp	sw
am	0.0	0.9	1.0	0.8	1.0	0.1	1.0	1.0	1.0	1.0	1.0	0.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0
ar	1.0	0.8	1.0	1.0	1.0	1.0	0.8	0.9	0.9	0.8	0.8	0.9	1.0	0.9	0.9	1.0	0.8	0.8	0.8
ch	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
da	0.9	0.9	1.0	0.2	0.8	0.9	0.9	0.8	0.9	1.0	0.9	0.7	0.8	0.8	1.0	0.6	1.0	0.8	0.9
du	0.9	0.9	1.0	0.7	0.7	0.9	0.8	1.0	0.9	0.9	0.9	1.0	1.0	1.0	0.4	0.9	1.0	1.0	0.9
en	0.3	1.0	1.0	1.0	0.8	0.1	1.0	1.0	1.0	0.9	1.0	1.0	1.0	1.0	0.9	0.9	1.0	1.0	1.0
fi	1.0	1.0	1.0	0.9	1.0	0.9	0.3	0.7	0.8	0.8	0.9	0.6	1.0	0.9	0.8	0.6	1.0	0.7	0.3
fr	1.0	0.9	1.0	0.9	0.8	1.0	1.0	0.2	0.8	0.9	0.7	0.9	1.0	0.9	1.0	0.8	1.0	0.9	0.7
ge	0.9	1.0	1.0	1.0	0.6	0.8	1.0	0.7	0.8	0.8	0.9	0.9	1.0	0.8	0.9	0.4	0.8	1.0	0.8
gr	0.8	0.9	1.0	0.6	0.9	0.6	0.8	0.3	1.0	0.5	0.7	0.9	0.8	0.7	0.2	0.8	0.8	0.7	0.7
hi	1.0	1.0	1.0	0.9	0.9	1.0	0.8	0.4	1.0	0.9	0.4	0.8	1.0	0.9	1.0	0.9	0.9	1.0	0.8
hu	1.0	1.0	1.0	0.8	0.9	1.0	0.9	0.6	0.9	0.9	1.0	0.8	0.8	0.8	0.9	0.8	0.7	0.9	0.8
it	1.0	1.0	1.0	1.0	1.0	0.9	1.0	0.9	1.0	1.0	0.9	1.0	0.5	1.0	0.9	0.9	0.7	0.9	1.0
no	0.9	0.9	1.0	0.7	0.8	1.0	0.4	0.6	0.9	0.9	0.8	0.7	1.0	0.2	0.4	0.8	0.8	0.8	0.8
pl	1.0	1.0	1.0	0.8	0.8	1.0	0.8	0.6	0.9	0.7	0.8	0.8	1.0	0.7	0.6	0.9	0.9	0.8	0.9
pr	1.0	1.0	1.0	0.8	0.8	1.0	1.0	0.7	0.8	0.4	0.9	0.9	1.0	0.9	1.0	0.5	1.0	1.0	1.0
ru	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.7	1.0	0.4	1.0	0.8	1.0	0.8	0.7	0.8	0.7	0.8	1.0
sp	0.9	1.0	1.0	1.0	0.9	1.0	0.8	0.8	1.0	0.9	1.0	0.9	0.8	0.6	0.9	1.0	1.0	0.7	0.6
sw	1.0	0.9	1.0	0.8	0.8	0.8	0.8	0.7	0.9	0.7	0.7	0.7	1.0	0.7	1.0	1.0	0.9	0.9	0.4

The following is the cluster analysis based on the confusion matrix by human listening test. This cluster analysis method is equal to the one in reference. The number of languages is 19, because Japanese is omitted. The above confusion matrix is not symmetric. Then we symmterize the table by averaging the two elements at i-th row, j-th column and i-th column, j-th row. In the table, the first column is the number of extracted clusters. The second column, distortion, is the average distortion in the case when each language is replaced by the representative languages, centroids, which is the third column. In the case of 7 language classes, AM, FR, CH, DA, FI, AR and NO are the representatives (centroids) of each language class. This result shows that AM and EN are close, CH and AR are isolated from the others, and also FI and SW are close. However, this result strongly depends on the listener's knowledge for languages.

No. of class	distortion	centroid
1	0.75	gr
3	0.62	am fr no
		am en
		fr ge gr hi it pr sw
		ar ch da du fi hu no pl ru sp
4	0.56	am fr ch no
5	0.52	am fr ch da no
		am en
		fr ge gr hi sw
		ch
		da du it pr
		ar fi hu no pl ru sp
6	0.50	am fr ch da fi no
7	0.49	am fr ch da fi ar no
		am en
		fr ge gr hi
		ch
		da du it pr
		fi sw
		ar
		hu no pl ru sp

H Choice of Files

american

am1f01	11.8	training
am1m01	12.0	training
am2f01	11.2	training
am2m01	9.9	training
am3f01	10.1	test
am3m01	9.3	test
am4f01	9.7	test
am4m01	10.5	test
am5f02	12.3	
am5f06	12.4	
am5m04	13.8	
am5m08	13.6	
am6f03	11.9	
am6f07	13.1	
am6m05	13.2	
am6m09	13.2	

arabic

ar1f01	5.4	training
ar1f02	3.7	training
ar1f03	4.2	training
ar1f04	4.7	training
ar1m09	3.5	training
ar1m10	2.9	training
ar1m11	5.6	training
ar1m12	6.5	training
ar2f05	8.2	test
ar2f06	6.0	test
ar2f07	4.8	test
ar2f08	3.8	test
ar2m13	3.4	test
ar2m14	5.4	test
ar2m15	5.5	test
ar2m16	4.0	test

chinese

ch1f02	7.2	training
ch1f10	5.5	training
ch1m01	6.0	training
ch1m09	6.0	training
ch2f04	5.7	training
ch2f12	6.0	training
ch2m03	5.5	training
ch2m11	4.1	training
ch3f06	8.1	test
ch3f14	5.0	test
ch3m05	5.5	test
ch3m13	7.2	test
ch4f08	6.9	test
ch4f16	5.3	test
ch4m07	6.2	test
ch4m15	7.8	test

danish

da1f01	13.0	training
da1f02	9.4	training
da1m01	12.6	training
da1m02	8.9	training
da2f01	14.4	training
da2f02	12.2	training
da2m01	11.6	training
da2m02	8.9	training
da3f01	13.4	test
da3f02	10.5	test
da3m01	12.1	test
da3m02	9.2	test
da4f01	13.4	test
da4f02	9.8	test
da4m01	9.5	test
da4m02	8.3	test

dutch

du1f09	8.7	training
du1f10	7.9	training
du1m01	8.7	training
du1m02	6.9	training
du2f11	5.9	training
du2f12	9.7	training
du2m03	9.0	training
du2m04	10.5	training
du3f13	10.3	test
du3f14	7.9	test
du3m05	5.3	test
du3m06	7.7	test
du4f15	9.0	test
du4f16	9.9	test
du4m07	8.5	test
du4m08	6.7	test

english

en1f01	6.3	training
en1m03	5.2	training
en2f02	6.4	training
en2m06	5.8	training
en3f04	6.6	training
en3m07	6.6	training
en4f05	6.6	training
en4m08	6.7	training
en5f11	6.6	test
en5m09	6.8	test
en6f12	5.7	test
en6m10	6.1	test
en7f14	5.7	test
en7m13	5.5	test
en8f15	6.9	test
en8m16	6.5	test

finnish

fi1f05	8.7	training
fi1f06	8.3	training
fi1m01	8.5	training
fi1m02	9.3	training
fi2f09	9.1	training
fi2f10	9.2	training
fi2m03	9.3	training
fi2m04	10.1	training
fi3f11	5.7	test
fi3f12	8.2	test
fi3m07	8.9	test
fi3m08	9.2	test
fi4f15	7.9	test
fi4f16	8.7	test
fi4m13	6.7	test
fi4m14	7.3	test

french

fr1f01	14.5	training
fr1f02	10.3	training
fr1m01	14.5	training
fr1m02	9.1	training
fr2f01	13.9	training
fr2f02	9.6	training
fr2m01	14.8	training
fr2m02	11.2	training
fr3f01	16.5	test
fr3f02	12.1	test
fr3m01	14.5	test
fr3m02	10.6	test
fr4f01	16.3	test
fr4f02	11.8	test
fr4m01	13.9	test
fr4m02	10.5	test

german

ge1f09	5.6	training
ge1f10	5.3	training
ge1m01	4.8	training
ge1m02	4.9	training
ge2f11	8.3	training
ge2f12	6.4	training
ge2m03	5.9	training
ge2m04	4.9	training
ge3f13	7.0	test
ge3f14	5.4	test
ge3m05	6.0	test
ge3m06	5.5	test
ge4f15	6.9	test
ge4f16	7.0	test
ge4m07	5.8	test
ge4m08	5.7	test

greek

gr1f09	7.3	training
gr1f10	9.2	training
gr1f11	8.7	training
gr1f12	10.3	training
gr1m01	7.7	training
gr1m02	8.9	training
gr1m03	8.8	training
gr1m04	7.4	training
gr2f13	10.2	test
gr2f14	9.1	test
gr2f15	11.8	test
gr2f16	10.6	test
gr2m05	8.9	test
gr2m06	12.3	test
gr2m07	11.4	test
gr2m08	8.0	test

hindi

hi1f07	7.9	training
hi1f08	5.5	training
hi1f09	6.3	training
hi1f10	7.4	training
hi1m01	7.5	training
hi1m02	8.4	training
hi1m03	8.6	training
hi2f11	7.6	test
hi2f12	8.0	test
hi2f13	6.1	test
hi2m04	5.1	test
hi2m05	6.2	test
hi2m06	6.7	test
hi3m14	6.1	
hi3m15	8.8	
hi3m16	8.5	

hungarian

hu1f03	6.9	training
hu1f14	10.8	training
hu1m01	7.0	training
hu1m02	7.3	training
hu2f06	4.8	training
hu2f16	7.5	training
hu2m04	6.2	training
hu2m05	5.2	training
hu3f07	4.6	test
hu3f08	7.5	test
hu3m11	9.3	test
hu3m15	8.9	test
hu4f09	10.5	test
hu4f10	9.6	test
hu4m12	9.1	test
hu4m13	8.7	test

italian

it1f09	12.3	training
it1f10	12.7	training
it1m01	7.3	training
it1m02	8.1	training
it2f11	9.0	training
it2f12	9.5	training
it2m03	8.9	training
it2m04	8.3	training
it3f13	7.6	test
it3f14	7.1	test
it3m05	9.4	test
it3m06	8.9	test
it4f15	7.7	test
it4f16	6.9	test
it4m07	7.9	test
it4m08	7.7	test

japanese

ja1f01	8.8	training
ja1f02	9.5	training
ja1m01	9.1	training
ja1m02	10.7	training
ja2f01	9.3	training
ja2f02	9.8	training
ja2m01	10.5	training
ja2m02	10.7	training
ja3f01	8.2	test
ja3f02	8.5	test
ja3m01	9.8	test
ja3m02	10.5	test
ja4f01	9.3	test
ja4f02	10.5	test
ja4m01	8.0	test
ja4m02	8.8	test

norwegian

no1f01	7.4	training
no1f02	7.2	training
no1m01	7.6	training
no1m02	8.2	training
no2f01	8.5	training
no2f02	9.7	training
no2m01	6.1	training
no2m02	6.0	training
no3f01	7.9	test
no3f02	9.0	test
no3m01	7.9	
no3m02	8.4	
no4f01	7.1	test
no4f02	7.9	test
no4m01	6.5	
no4m02	7.1	
no5f01	5.7	test
no5f02	6.8	
no5m01	5.4	test
no5m02	5.9	test
no6f01	6.7	test
no6f02	7.0	

polish

pl1f01	7.2	training
pl1f02	11.7	training
pl1m03	9.9	training
pl1m04	9.5	training
pl2f07	5.8	training
pl2f08	9.4	training
pl2m05	6.8	training
pl2m06	7.4	training
pl3f11	8.7	test
pl3f12	10.3	test
pl3m09	10.7	test
pl3m10	9.8	test
pl4f13	8.1	test
pl4f14	7.7	test
pl4m15	5.2	test
pl4m16	8.4	test

portuguese

pr1f01	6.4	training
pr1f02	10.4	training
pr1m03	5.5	training
pr1m04	9.4	training
pr2f07	7.6	training
pr2f08	6.7	training
pr2m05	6.5	training
pr2m06	5.8	training
pr3f13	6.5	test
pr3f14	7.8	test
pr3m09	8.0	test
pr3m10	6.5	test
pr4f15	6.6	test
pr4f16	8.7	test
pr4m11	5.5	test
pr4m12	8.2	test

russian

ru1f09	9.0	training
ru1f10	11.1	training
ru1m01	10.5	training
ru1m02	10.1	training
ru2f11	6.8	training
ru2f12	9.4	training
ru2m03	6.7	training
ru2m04	4.9	training
ru3f13	6.1	training
ru3f14	7.5	test
ru3m05	9.8	test
ru3m06	5.6	test
ru4f15	7.1	test
ru4f16	2.2	test
ru4m07	6.8	test
ru4m08	7.6	test

spanish

sp1f03	6.6	training
sp1f04	10.5	training
sp1m01	8.1	training
sp1m02	5.8	training
sp2f07	7.3	training
sp2f08	4.3	training
sp2m05	8.0	training
sp2m06	9.1	training
sp3f11	7.0	test
sp3f12	7.2	test
sp3m09	7.2	test
sp3m10	7.4	test
sp4f15	6.5	test
sp4f16	5.7	test
sp4m13	10.1	test
sp4m14	9.3	test

swedish

sw1f01	7.4	training
sw1f02	7.6	training
sw1m09	6.6	training
sw1m10	6.3	training
sw2f03	7.9	training
sw2f04	7.1	training
sw2m11	7.1	training
sw2m12	6.0	training
sw3f05	6.9	test
sw3f06	6.2	test
sw3m13	7.7	test
sw3m14	7.0	test
sw4f07	7.3	test
sw4f08	6.7	test
sw4m15	8.0	test
sw4m16	7.2	test

I Sentence lists in 20 Languages

american	
am1f01	The ship was torn apart on the sharp reef. Sickness kept him home the third week. The box will hold seven gifts at once. Jazz and swing fans like fast music.
am1m01	
am2f01	
am2m01	
am3f01	
am3m01	
am4f01	
am4m01	
am5f02	A rod is used to catch pink salmon. The source of the huge river is the clear spring. Kick the ball straight and follow through. Help the woman get back to her feet.
am5f06	High seats are best for football fans. Tea served from the brown jug in tasty. A dash of pepper spoils beef stew. A zestful food is the hot-cross bun.
am5m04	The birch canoe slid on the smooth planks. Glue the sheet to the dark blue background. It's easy to tell the depth of a well. Four hours of steady work faced us.
am5m08	The swan dive was far short of perfect. The beauty of the view stunned the young boy. Two blue fish swam in the tank. Her purse was full of useless trash.
am6f03	A siege will crack the strong defense. Grape juice and water mix well. Roads are paved with sticky tar. Fake stones shine but cost little.
am6f07	Heave the line over the port side. A lathe cuts and trims any wood. It's a dense drowd in two distinct ways. His hip struck the knee of the next player.
am6m05	Rice is often scrved in round bowls. The young kid jumped the rusty gate. Guess the resultes from the first scores. A salt pickle tastes fine with ham.
am6m09	A speedy man can beat this track mark. He broke a new shoelace that day. The coffee stand is too high for the couch. The urge to write short stories is rare.

arabic	
ar1f01	tutabaru lgam iyyatu lammatu gihaza lmuttahidati rraisiyya li igrai lmudawalati
ar1f02	yaqau maqaru munazzamati lumami lmuttahidati fi nyuyrk
ar1f03	rabbi innaka wahabtani nagahan fala tanza tawadui
ar1f04	raaytu nnafsa takrahu ma ladayha wa tatlubu kulla mumtaniin alayha
ar1m09	wa qul rabbi rhamhumma kama rabbayani sagira
ar1m10	ala qadri ahli lazmi tati lazaimu
ar1m11	kulu man talqahu yasku dahrahu layta siri haihi ddunya liman
ar1m12	al umamu lmuttahidatu hiya itihadun munazzamun bayna duwwalin mustaqillatin ati siyyadatin
ar2f05	wa ahiran ittafaqa ladau ala irgai ligtimai ila agalin gayri musamma
ar2f06	laysa lbirru an tuwallu wughakum gibala lmasriqi wa imagribi
ar2f07	
ar2f08	dahala rragulu lbayta wa lam yahrug minhu
ar2m13	awaddu awwalan wa qabla kulli say in an urahhiba bikum gamian
ar2m14	la yahfa alaykum anna lhadafamina qtimaina huwa tabadulu larai bisa ni ssilmi lalamiyyi
ar2m15	yuaddi lmutarqimuna lfawriyyuna dawran hamman fil idtilai bi muhimmati littisali ssaqati
ar2m16	qunit madinatun raiatun lakin tanqusuha rrimalu ddafiatu

chinese		
ch1f02		
ch1f10		
ch1m01		
ch1m09		
ch2f04		
ch2f12		
ch2m03		
ch2m11		
ch3f06		
ch3f14		
ch3m05		
ch3m13		
ch4f08		
ch4f16		
ch4m07		
ch4m15		

danish	
da1f01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da1f02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da1m01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 xeller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da1m02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han a ltid.
da2f01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da2f02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da2m01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da2m02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da3f01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da3f02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da3m01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da3m02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da4f01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da4f02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.
da4m01	Det er ikke s@rlig sv@rt at finde Rigshospitalet. Du star bare af ved Osterport station, og sa tager du linie 1 eller 6 til Trianglen, og linie 3 videre et stykke ned ad Blegdamsvej.Sa kommer det pa hOjre side - det er vist allerede det andet stoppested.
da4m02	Jeg sadge til Peter at det ville jeg altsa ikke finde mig i.Og ved du hvad han svarede? Han sagde:"Hov-hov! Ikke noget med at komme og spille fr@k over for migi!"Er det ikke bare utroligt? Sadan er han altid.

dutch	
du1f09	Nu bevond zich in ons nidden een psychiater, een scherpzinnig en oppassend man, die het verschihnsel met groeiende verbazing gadesloeg.
du1f10	Hierdoor opmerkzaam geworden bestelden we opineuw thee en zetten ons aan het raam, om het verschijnsel eens nauwkeurig op te nemen.
du1m01	Postkoetsen hebben de meeste kans overvallen te worden als de passagiers bestaan uit een dominee, een valsspeler, een meisje van lichte zeden en een held.
du1m02	Het scherpe proza van deze schrijver sttat in schril contrast tot de onbeholpen wijze waarop hij zich mondeling uitdrukt.
du2f11	Het is moeilijk een discussie te winnen wanneer je tegenstander bezwaard is met een goede kennis van feiten.
du2f12	De aanleg van de baan kost weliswaar tien miljoen, maar de stichting die de baan moet gaan beheren denkt dit bedrag voor een groot deel uit subsidies bijeen te kunnen krijgen.
du2m03	Hij blijft ernnar turen, net zo lang tot het hobbelen opeens verdwijnt en hij onder zijn voeten het bonkende intrekken van het landingsgestel voelt.
du2m04	Het zwaarst getroffen zijn de binnenstad, waar naar schatting zestig procent van de gebouwen is beschadigd, en de krottenwijken, waar de meeste huizen beschadigd of vernietigd zijn.
du3f13	We staan niet aan het einde van een oude periode, we staan aan de deur van een geheel nieuw tijdperk en we hebben nog niet eens gebeld om erin te mogen.
du3f14	In een mengeling van welbehagen en spannnig laat Arnold zijn in de rugleuning van zijn stoel drukken en kijkt naar buiten.
du3m05	Alles, wat uitsteekt,wekt in de Nederlander de onweerstaanbare drang om het gelijk te maken met zijn omgeving
du3m06	In de bovenverdieping van het hotel waren zes frisse, keurige slaapkamers voor de gasten, en beneden twee zalen, een grotere en een kleinere.
du4f15	Als ik veertien dagen op een hotelkamer ben, dan is die hele kamer dichtgeslibt met de kruimels die de stad eromheen te bieden heeft.
du4f16	Van de gelegenheid tot het stellen van vragen zag ieder grootmoedig af, zeer tot ongenoegen van de voorzitter die nog maar kort in het bestuur was.
du4m07	Er zijn twee soorten vuil: het donkere dat aangetrokken wordt dor lichte voorwerpen, en het lichte dat aangetrolken wordt door donkere voorwerpen.
du4m08	Onder zijn hoge schoenen kraakt de sneeuw, die het licht van de gaslantarens omzet in een droomachtig schijnsel.

english	
en1f01	I was away for nine weeks. The dining-room was lit by gas. There were no vegetables left.
en1m03	I think it would be wonderful. There will be guests coming. She deldom listened to anybody.
en2f02	It's human nature to blame another.He had completely forgotten his hat. I did not wish him to know.
en2m06	He could not remember his name.I never can leave you two alone. I shall entreat his pardon.
en3f04	The act was a deliberate murder. He was attracted by her face. The timber fell across the road.
en3m07	They have a good idea of fairness. You must go and do it at once. Have you said all you have to say?
en4f05	We were dompelled to stay there. He had a curious sence of power. This car appears to have six wheels.
en4m08	He was not in the mood for music. I can well understand your feelings. I have rented a small house.
en5f11	This is no business of yours. What did he say he went up for? She was in bed when she got home.
en5m09	Not a flower bloomed in the garden. I have never had time to learn. She was not impressed by his talk.
en6f12	The king had a navy at sea. I have told you everything. Have you the letter with you?
en6m10	We poured cold water on his head. There was not much time left. It is useless wating for a taxi.
en7f14	He came running up to us. You haven't got the words right. My cup of trouble is now full.
en7m13	He looked about him again. They met agian in the station. Why must there be all this secrecy?
en8f15	The storewas open on Saturday.She always asked him what to do. Now they had some packing to do.
en8m16	I saw it with my own eyes. The note was immediately dispatched. He wanted to leave college.

finnish	
fi1f05	Useimmat alueen tunturit nostavat lakense hadin tuskin puurajan ylapuolelle, ja usein laen puuttomuus johtuu pikemminkin kivisyydesta kuin korkeudesta.
fi1f06	Etelampana ja suotuisemmissa oloissa opittuja metsatalouden menetelmia on kaavamaisesti sovellettu myos pohjoisilla metsanlasvun aarirajoilla.
fi1m01	Undistusyritykset ovat koskeneet hyvin laajoja alueita, ja nykyaan on monin pailoin vain rippeita jaljella ennen niin yleisesta paksusammaltyypista
fi1m02	Nykyaika kelloineen, radioineen ja kiinteine tyoaikoineen on paljon rajoittanut vapaata ajankayttoa, mutta silti moni lahtee tyohonsa vasta illan viilennyttya.
fi2f09	Kevaalla puiden ja pensaiden silmut ja syksyn kuihtunut aluskasvill suus varjaavat nakymaa hennoin savyin, joita muulloin tuskin tapaa.
fi2f10	Viljelykelpoinen maa on nailla, maamme ilmastollisesti parhaimmilla alueilla, pyritty ottamaan mahdollisimman tarkkaan voimaperaiseen viljelyyn.
fi2m03	Maankohoaminen on muutellut kaikkien Suomen suurten jarvien aariviivoja, ja hyvin usein on myos laskukynnys siirtynyt toiseen jarveen.
fi2m04	Sisasaaristo on nykyaikaa, kun taas ulkosaaristo kertoo, millainen sisasaaristo on ollut vuosituhansia sitten saarten vasta paljastuttua meresta.
fi3f11	Ymmarrettavasti talven rajaksi on valittu ajankohta, jolloin vuorokauden keskilampotila alittaa nolla-asteen.
fi3f12	Viela pitkalle talveen vesi solisee lammista jarviin ja jarvista mereen, ja naissa sulapaikoissa elama sykkii nakyvammin kuin lumivaipan suojassa.
fi3m07	Luonnonkauniiden saaristonakymien ohella matkailijalle on tarjolla vehmaita kulttuuriseutuja vauraine maatalousalueineen ja historiallisine rakennuksineen.
fi3m08	Entisointityossa pyrittiin linnan tilat samaan alkuperaisenn asuunsa niin, otta osa niista toimisi museotiloina ja osa olisi paivittaisessa kaytossa.
fi4f15	Laitumella kayvat elaimet estavat puiden taimien kasvun, minka johdosta puut ovat lehtoniityilla puistomaisen harvassa.
fi4f16	Mutta tanaan otettu valokuva voi jo aikaisemmin kkuin sadan vuoden paasta, samoilta jalansijoilta maisemaan verrattuna, nayttaa oudolta.
fi4m13	Poronhoidon vaatimukset huomioon ottaza ratkaisu petokysymyksessa onkin puiston tulevaisuuden kannalta keskeinen seikka.
fi4m14	Alarinteiden metsat ovat enimmakseen karuja mantykankaita, jotka kansallispuistossa ovat saaneet olla hakkuutoiminnalta rauhassa.

french	
fr1f01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr1f02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr1m01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr1m02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr2f01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr2f02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr2m01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr2m02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr3f01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr3f02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr3m01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr3m02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr4f01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr4f02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.
fr4m01	La bise et le soleil se desputaient,chacun assurait qu'il etait le plus fort, quand ils virent un voyageur s'avancer enveloppe dans son manteau. Ils tomberent d'accord que celui qui arriverait le premier a lui faire enlever son manteau serait le plus fort.
fr4m02	Alors la bise se mit a souffler de toutes ses forces ; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et a la fin la bise renonca a le lui faire enlever.

german	
ge1f09	Ich vermute, da@ die Leute auf ihrer Autofahrt keine haufigen Probleme mit den hangenden Sonnenuhren hatten.
ge1f10	Leider gibt es bei Dir zu Haus wegen der Teuerung erst am dienstag wieder frisch gebrannte Mandeln.
ge1m01	Sieh an ob da nur einer der Steine liegt, die ich im winter anderwarts gefunden habe?
ge1m02	Geld allein macht niemanden glucklich, aber verdienen nicht bose Menschen ihre Strafe?
ge2f11	Jawohl, horen Sie! Ich bin Ruldolf Ranick hier vom FTZ. Prufen Sie bei Kurt Meier in der Burgstra@e den Leitungsanschlu@!
ge2f12	Nervose Menschen brauchen namlich viel Ruhe, darum kostet heute jeder Platz durch di Bank eins funfzig.
ge2m03	Ich unterstelle,da@ die Bruder wahrend der Reise ja auch haufig eng unter Kokospalmen gesessen haben.
ge2m04	Mittwoch kommt uns der Besuch ja passend, denn dann bin ich noch nicht wieder na@ geworden.
ge3f13	Das Preisgericht hat nun endlich entschieden: Zuviel Syntax qualt ja doch gewi@ die murben Kopfe.
ge3f14	Am kuhlen Abend wird unser nachster Treffpunkt wohl wieder gegen zwei Uhr am Neumarkt sein.
ge3m05	Die Bruder Ingo und Ivo Kaiser durften taglich unter schlanken Pinien jede Menge Bier getrunken haben.
ge3m06	Larmt nicht Jungs, denn Vater schreibt noch, aber nachher tanzen unsere Eltern Wiener Walzer.
ge4f15	Nach dem Halali sa@en in der Nacht schlie@lich zwolf Jager quasi gebuckt vor dem Xylophon.
ge4f16	Ursel weint zwar, aber Heinz lacht ganz leise, weil Adokl nun wohl auch Lehrer werden will.
ge4m07	Du Kannst ja allmahlich mit den Jahren auch langst nicht mehr alle hohen Baume gabz schnell erklettern!
ge4m08	Wei@t Du denn nun genau Bescheid, wen Du dort auf welcher Seite links oder rechts gehen siehst?

greek	
gr1f09	O yatros too ipe oti prepi na kani thieta ya na hasi pende kila ke na min kapnizi oote na pini inopnevmatothi.
gr1f10	Prin erthis apopse spiti, perase se parakalo apo to katharistirio na paris ta sendonia ke tis maxilarothikes poo eho thosi ya plisimo ke sitheroma.
gr1f11	Fevgo stis thekaexi Oktovrioo ke tha lipso ya tesseris evthomathes yafto then tha boreso na se tho prin apo tis thekapende noemvrioo.
gr1f12	Tote moo irthe sto mialo oti xekinisa thihos ligo psomotiro thihos ena pagoori nero ya na anevo se tootes tis erimies se toota ta afiloxena vrahia.
gr1m01	I athekfi soo perase hthes to vrathi apo to spiti na me thi ke moofere to vivlio ya to apio milisame to perasmeno savato. Tha soo to epistrepso molis to thiavaso.
gr1m02	Htbes to aoiyevma pire to leoforio ke katevike sta mayazia me tin yati ithele na agorasi ena poolover ala then vrike tipote na tis aresi.
gr1m03	Afoo perpatisame kamia ora, frasame sena mikro estiatorio konda sto potami opoo boresame na pyoome kati na xethipsasome ke na fame mesimeriano.
gr1m04	I fill moo i Maria ine poli stenohorimeni afton to kero, ehi pola provlimata me ton yo tis, poo then ta pai katholoo kala sto sholio.
gr2f13	Otan itane nea piyene sihna to kalokeri tin ora too thilinoo ke kathotan sto limani ke zografize tis psarovarkes ke ta kaikia.
gr2f14	Tha ithela na soo po na min me paris tilefono apopse yati tha lipo apo to spiti, tha pao ston kinimatografo me kati filoos.
gr2f15	Pira ena kafe ke ospoo na ton kalopio eftase ke to leoforio ke stathike isia bros sto kafenio.Anevika ke vrika mia thesi thipla ston othigo.
gr2f16	I Klio afoo teliose to yimnasio me atista piye stin Aglia na spoothasi viohimia ekane ke tin thiatrivi tis ke tora vriskete stin Ameriki opoo ergazete.
gr2m05	Moo ipe oti molis boresi na mazepsi ta apetoomena hrimata tha kani ena megalo taksithi sti notio Ameriki mazi me ti yineka too ke ta pethia too.
gr2m06	Ya ta egenia too fragmatos eyine episimi teleti itan o Prothipoorgos o Ipoorgos Thimosion Ergon itan ke o Mitropolitis ke evloyise ke o thimarhos tis polis evgale logo.
gr2m07	Otan i Yermani tin toofekisan stis arhes too hilia eneakossia saranda tessera, ligo prin apo tin apeleftherosi, i Margarita then ihe patisi akoma ta ikosi hronia tis.
gr2m08	O papasn ipe pos otan sikosan ta doofekia, i mikri Margarita koonise to heri tis ke ipe ena akatanoito kalinihta.

hindi	
hi1f07	Tumhai naukari mai rakhanai sai pahalai so bar socha mainai tumhai rakhu ya na rakhu na rakhtha tho bat agai nahi badathi.
hi1f08	Gharvalon ki jali kati sonthi rahi par mere man mai kabhi nahi aya ki mai tyag kar rahee hoon.
hi1f09	Aap chalai aayaihai uskee vakalat karanee par kabhi apanee bhumika kai bari mai socha aapanai.
hi1f10	Mishra eesa shahar mai jab tak rahaiga aap aapnai kandho sai bithi hoee ghtanaoon kee lasha nahi utar sakega.
hi1m01	Mai mantha hu ke daphtar ki phone ka estimal kival daftar ki kam ki leya he hona chahiya niji kam ki leya nahe.
hi1m02	Yahi ke agar bade babu ki sali ke niukati aapki jagah ho jathi tho phalai hi din us par itana kam ladha diya jata kya.
hi1m03	Abdul shadi shuda hi nahi do bacho ka bap hai rat ko baitha-baitha chithiya likha karatha hai apani biwi ko.
hi2f11	Tunai uski sonai ki mala or chudiyan baichakar sudeep ko paisai dai diya laikin usanai kucha nahi kaha.
hi2f12	Phichalai panca salao mai ek-ek paisa dant sai pakadkar rakhatha kya isliya ke tumhara bhai udan bhari or hamai jameen dikha dai.
hi2f13	Kshan tho mratui,kai ketanai karib pahuncha gaya tha or kahan rumnai mughi jeevan sai eetanai gaharai todha diha.
hi2m04	Hamari samasya hai garibi kai raigistan mai mahatvakakshao ka kamal khilani ki jedha.
hi2m05	Ek bat yad rakhana chaya samai kai hat kamai aam ki thara chuskar a likhit itihas kai dhurai par phika daigai.
hi2m06	Ab hamari chintan-shakti samapt ho chuki hai sochathi nahi isaliyai zinda hai sochaingai tho mar jayaigai.
hi3m14	Mai madhamyargiyai naithikata ko lagatar tumharai samnai dhonga kahata raha par tum apanee jagaha atal rahi.
hi3m15	Jab modi nai tumhari samnai shadi ka prastav rakha usi samaya mujnai virodh karna tha jakar sidhai modi sai lada padana tha usi lalkarna tha.
hi3m16	Mujhaye samjha mai aa raha hai ke jin prashanon par aamanai aakar ladaee karana jaroori hai jinkai phisalai sadakoon par kiya janai chahiai.

hungarian	
hu1f03	A homersekleti ingadozas miatt a keltetogepek nem teljesithetik jol a tervezett atlagokat.
hu1f14	Hetente tobb alkalommal viszik a tanulokat az uszodaba, megis igen alacsony azoknak a szama, akik megbizhatoan tudnak mar uszni.
hu1m01	A falatozoban sort, bort, uditoitalokat es finom malacsultet lehet kapni.
hu1m02	Az ezredes felesege banatosan abrandozott, miutan a piros pecsetes levelet elolvasta.
hu2f06	Nehany egyetemi es foiskolai karon iden megint sokszoros volt a tuljelentkezes.
hu2f16	A vedett madarak erteket annak alapjan becsultek fel, hogy az egyes videkeken milyen mennyisegben fordulnak elo.
hu2m04	A hadsereg tisztjeit meg aznap kivezenyeltek, hogy tudassak veluk a legujabb napiparancsokat.
hu2m05	Az altalanos iskolakban ebben az evben is szeptember elsejen kezdodott a tanitas.
hu3f07	- Mit gondolsz jo ez a szindarab? - Melyik? - Amelyiket a mult heten lattunk.
hu3f08	Nyaranta alig-alig kopogtatnak a korzetorvosi rendelok ajtajan, ugyanakkor osszel es telen rendkivuli a tolongas.
hu3m11	Az ures uvegeket kosarban vittuk a boltba, de legnagyobb meglepetesunkre az uzlet mar negy orakor zarva volt.
hu3m15	A porcelan hosszu evszazadokig Azsia csodalt termeke maradt, hasztalanul kiserleteztek, Europaban nem sikeruly hasonlo anyagot gyartani.
hu4f09	Az ovodak egy rezebol iskolakat hoztak letre, mivel a rendelkezesre allo tantermek nem voltak elegendoek az elso osztalyosok fogadasara.
hu4f10	Regota tudjuk, hogy mindenki ele olyan celt kell kituzni, amit feltehetoen el tud erni, ekkor remelheto a sikerelmenye.
hu4m12	- Matyikam, miert veted le a cipodet? - Mert belement egy kavics. - Es miert veted le mind a kettot? - Hogy lassam, melyikbe ment bele.
hu4m13	Szerfelett grotteszknek tuno otlet 19444-ben egy magyar kisvarosban attol felni, hogy az embereknek agyukra megy a jolet.

italian	
it1f09	La valle e circondata dai montti La terra gira attorno al sole Il coniglio mangia l'erba Le suole delle scarpe sono di cuoio
it1f10	Gli occhiali servono per vedere Il canguro vive in Australia La motocicletta ha due ruote La Spagna confina con la Francia
it1m01	Le rose fioriscono in maggio L'India e un paese misterioso Il tiume ha rotto gli argini Lo scolaro e stato promosso
it1m02	Le strade liguri sono pittoresche Il mare e molto agitato Il gatto e un animale agile La siepe separa il campo arato
it2f11	La tua valigia e di pelle nera @opo il temporale viene sereno Il sangue coagula rapidamente La benzina s'incendia subito
it2f12	Il vetro si rompe facilmente La freccia ha colpito il bersaglio La giungla e infestata dai serpenti I cittadini pagano le tasse
it2m03	I sacchi sono molto pesanti La sabbia del deserto e rovente Il dente di Paolo era cariato Il bisturi era molto tagliente
it2m04	Quel bambino ha i capelli rossi Ieri il medico era di guardia Le nubi coprono le alte cime Il dolore era molto intenso
it3f13	La Iavandaia sciacqua i panni L'areoplano vola molto alto La barca e ormeggiata al molo Dopo il lavoro occorre nposo
it3f14	Le onde arrivano sulla spiaggia Il cannone ha sparato un colpo La moto e uscita fuori strada Il compito era molto difficile
it3m05	Le acque del mare sono agitate Il chirurgo era molto stanco Il vaso cade dalla finestra La bandiera d'Italia e tricolore
it3m06	Stasera il cinema e deserto La paziente era molto agitata Il cancello era chiuso a chiave Gustavo e caduto dalle scale
it4f15	Una nuova guerra minaccia il mondo Le liane sono molto robuste Il poscatore e senza licenza Il dattero e un frutto prelibato
it4f16	Il motoscafo e molto veloce Il Polesine e spesso allagato Il gioco delle bocce e divertente Il delfino e molto socievole
it4m07	La bonta e una dote preziosa La garza tampota la ferita Il colpepvole e stato condannato Girare il mondo e divertente
it4m08	Lo specchio riflette le immagini Il tribunale condanna i ladri La cascata fa molto rumore I topi rosicchiano il formaggio

japanese	
ja1f01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja1f02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja1m01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja1m02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja2f01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja2f02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja2m01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja2m02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja3f01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja3f02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja3m01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja3m02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja4f01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja4f02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru
ja4m01	karewaizen'kara/kagakugijyutunosin'potonin'gen'noyoukiga/ harukanautyuenotabio/kanounisitanodatokan'gaeteimasita
ja4m02	hitobitono/byoubuetonyoraizounitaisurukyoumiwa/ happyakunen'nonen'getuniyottesyoujita/hyoumen'no bimyounasikisaihen'kaniaru

norwegian	
no1f01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no1f02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no1m01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no1m02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no2f01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no2f02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no2m01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no2m02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no3f01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no3f02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no3m01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no3m02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no4f01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no4f02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no4m01	Sjefen har benyttet datamaskin som viktig hjelpemiddel i jakten pa gunstige bolgeform- og kildekodere.
no4m02	I dette norske bidraget til kunstig talesignal er kvinner og menn, men ingen forskolebarn, med som opplesere.
no5f01	
no5f02	
no5m01	
no5m02	
no6f01	
no6f02	

polish	
pl1f01	Straz miejska, po dlugich poszukiwaniach sprawcy mordu, dzieki przypadkowi wykryla go i ujela bez oporu.
pl1f02	Wlodzimierz Potocki z gniewem bezsilnym pogladal na piekny swoj oddzial, na nowiutkie mundury kanonierow, na blyszczace dielska armat, na nie zbrokane jeszcze lawety.
pl1m03	Szesc koni oprzegalo kazda, armate - koni roslych, o szerokich piersiach, o kartkach krotkich, miesistych i nosach grubych, wlochatych.
pl1m04	Chirurg nie dal sie zwatpieniu i zabral sie z kolei do opatrywania rozwartej piersi pulkownika, gdzie krwotok byl nierownie mniejszy niz na ramieniu.
pl2f07	Jego cesarska i apostolska mosc zywi ku wam najserdeczniejsze intencje.
pl2f08	Co chwila ktos zapalczywy goretsze rzucal slowo, co chwila trzaskaly zlowrogo palasze a z ocz przeciwnikow padaly skry.
pl2m05	Koszary Ujazdowskie wypelniono po brzegi, stajnie Mirowskie zawalono cialami oprozniono domy i palace.
pl2m06	Chocby teraz wlasnie nadarza sie sposobnosc, abyscie laczac sie z nami, zjednali sobie trwalego i wiernego sprzymierzenca.
pl3f11	Prosze pana, pewna kwoka traktowala swiat z wysoka i mowila z przikonaiem:"grust to dobre wychowanie".
pl3f12	Przy rozpatrywaniu tworczosci Szopena, pierwszoplanowym i wezlowym problemem staje sie, sprawa jego stylu oraz wartosci i ich wzajemnego wspokgrania.
pl3m09	Widok rozmawia hacych w oddali wodzow przywracal spokoj, budzac znow pozorna harmonie, aby przy nowym nieporozumieniu jeszcze gwaltowniejsza wywolac wrzawe.
pl3m10	Bylo sobie raz jajko madrzejsze od kury . Kura wylazi ze skory, prosi, blaga, namawia:"badz glupsze " - lecz co poradzic, kiedy ktos sie uprze?
pl4f13	Placze chrzan na salaterce, az sie wszustkim kraje serce:"Panie chrzanie niech pan przestanie".
pl4f14	Tworczosc nuzyczna jest zroznicowana pod wzkledem gatunkow i form oraz prechodzi wyrazna ewolucje.
pl4m15	Najblizsze lata po kongresie wiedenskim nie byly latwe dla szwajcarskiego zwiazku panstw stowarzyszonych.
pl4m16	Sqrawa ktora wstrzasnela opinia publiczna i doprowadzila do powaznych zmian wewnetrzny w systemie rzadow konstyutucy jnych, bylaslawetna sprawa samolotow szpiegowskich.

portuguese	
pr1f01	No mesmo e agradavel espaco, escolinhas de esporte fazem o lazer dos fins de semana.
pr1f02	A partir dos 26 anos de idade - tempo de vida de uma pessoa extremamente jovem para os atuais padroes - a mulher vai perdendo o colageno e a elastina.
pr1m03	Vamos economizar a natureza, pois ela e o combustivel da vidaOC.
pr1m04	Cerca de 300 alunos do Colegio Sao Vicente de Paula, em Icarai, assistiram a uma palestra sobre Telefonia.
pr2f07	Houve um incendio de pequenas proporcoes na torre de arrefecimen to da estacao Leblon localizada na Rua Humberto de Campos.
pr2f08	O filho de ZEUS concIderado no Olimpo, bebia o nectar e alimentava-se com ambrosia.
pr2m05	Os Distritos de Macae e Cabo Frio estao juntos no mais novo distrito da TELERJ.
pr2m06	A Diretoria Tecnica pretende ampliar o numero de Kombis para tes te de propagacao.
pr3f13	Em certo sentido o pai que sacrifica o filho acaba sendo condenado a suplicios.
pr3f14	Fabio Barreto vai interpretar um solteirao conservador que acaba nao percebendo as mudancas no pais.
pr3m09	Evite bloqueios da expressao emocional que se acentuam com Venus retrogrado - fase de intensa sensualidade.
pr3m10	Ha portanto, igualmente um potente conteudo invejoso no ato sacrificial de TANTALO.
pr4f15	O casal alias acabou driblando os fotografos, pois chegou em cima da hora e saiu antes do fim.
pr4f16	As tres safras de graos por ano sao: arroz, trigo e num pomar de laranja, batata, soja, jasmim e amendoim.
pr4m11	Os novos veiculos sao equipados com mastro telescopico elevado por sistema pneumatico.
pr4m12	Comparado ao metodo convencional o novo teste de propagacao ofere ce inumeras vantagens, diz o Cheie da Dividao de Meios de Transmissao.

russian	
ru1f09	
ru1f10	
ru1m01	
ru1m02	
ru2f11	
ru2f12	
ru2m03	
ru2m04	
ru3f13	
ru3f14	
ru3m05	
ru3m06	
ru4f15	
ru4f16	
ru4m07	
ru4m08	

spanish	
sp1f03	Un jinete se separo de la sombra que formaban los arboles junto a la ca rretera y se dirigio lentamente hacia la choza.
sp1f04	En aquella llanuta habia 12 rebanos con un total de 25 carneros, 4.763 ovejas y 1982 corderos lechales.
sp1m01	Bajo el veello de una barba de una semana, se distinguia confusamente el perfil de la mandibula y del menton, y habia manchas oscuras debajo de sus ojos.
sp1m02	Esa senora venia mucho a mi casa y, a veces, me ayudaba a blanquear y limpiar los armarios de la cocina.
sp2f07	Despues de unos 340 dias de gestacion, la yegua da a luz a un potro que es capaz de seguir a su madre al cabo de muy poco tiempo.
sp2f08	Joaquin no se daba cuenta del gran papel que su hijo desempenaba en aquellos momentos.
sp2m05	La Ecologia es la rama de la Biologia que trata de las relaciones entre los seres vivos y el medio ambiente.
sp2m06	En la fotosintesis, los vegetales toman la energia de la luz mediante la clorofila para formar su materia organica.
sp3f11	Dentro del baul y en bolsas de celofan, tenia queso manchego, vino ane jo de la tierra y botellas de leche.
sp3f12	La ciguena es una ave zancuda, con pico y patas muy largas, que se alimenta de pequenos vertebrados e insectos.
sp3m09	En varias de las cabanas unos ninos de pecho rompieron a llorar con unos chillidos tales que nada parecia poder parar.
sp3m10	La habitacion da a una plaza antigua en la que se levanta un edificio grande y rectilineo rodeado de casas bajas.
sp4f15	Se llaman fosiles a los restos o huellas de animales o plantas que exis tieron antiguamente y se conservan petrificados.
sp4f16	La flor es el organo reproductor de las plantas fanerogamas y de ellas se originan los frutos y las semillas.
sp4m13	La gangrena es la muerte local de tejidos por falta de oxigeno, producida por causas fisicas, quimicas, circulatorias, nerviosas o infeccio sas.
sp4m14	El elefante es el mayor de los mamiferos terrestres y dispone de una trompa que es el resultado de la prolongación de a nariz y del labio superior.

swedish	
sw1f01	Skorna var nya och alldeles for tranga men lingonen brukar mogna i september fast pumpen pa garden hade rostat pa vintern.
sw1f02	Applet ar moget och alldeles gront fast bussens forare fick korkortet indraget eftersom isen omojligt kan bara en vuxen.
sw1m09	Kajsa grat nar ballongen sprack och ett vitt sken syntes vid horisonten nar samtalet stordes av bullret pa gatan.
sw1m10	Klipporna vid kusten ar hala och graa men hamnen doftar tjara och farg och trad och buskar gronskar redan.
sw2f03	Vagorna slog hogt over bryggan i stormen da dagen firades med klang och jubel trots att snogubben har tappat fyra knappar.
sw2f04	Torpet hade blommor och gras pa taket men manga trivs med att vandra i fjallen trots att sikten ar ganska skymd i kurvan.
sw2m11	Lamporna lyser knappast av sig sjala utan av kampanjen som avslutas redan pa sondag da solen har blekt mina basta gardiner.
sw2m12	Ulla blev instangd nar dorren stangdes och Eva har tappat nyckeln till cykeln fast Kerstin har klamt sin vanstra rumme.
sw3f05	Gunnar salde en korg med blommor i Jonkoping som ligger vid Vatterns sydspets och yxan maste slipas for att falla tradet.
sw3f06	Batar och bryggor hor sommaren till men det harliga lovet ar nastan slut och snart har den nya terminen borjat.
sw3m13	Tradgardsarbete ger harda och valkiga hander men resultaten av proven ar helt perfekta fast han paddlade kanot runt hela kusten.
sw3m14	Bjorkarna skiftar i lila pa varen och snodropparna blommar redan i Skane medan goken gol nagonstans i vaster.
sw4f07	Kalle laste sina laxor slarvigt om hur hunden och katten leker tillsammans fast gungorna i parken ar alldeles for hoga.
sw4f08	Skramla inte med koppar och fat fast golvet ar nybonat och valdigt halt sa spetsa oronen och lyssna noga.
sw4m15	Takpannorna blaste ner i den svara stormen och kattungen rev sonder de nya gardinerna for flickan som har roda rosetter i flatorna.
sw4m16	Det ar alldeles for varmt att sitta i buss och turisterna finns pa alla badorter medan alla teatrar har stangt pa sommaren.

J Figures

J.1 Experimental Results

J.2 Sonagram for several sentences

J.3 Spectra in generated codebook

J.1 Experimental Results

parameter	value
distance	WLR
amount of test data	1
frame shift for training	128
frame shift for test	512

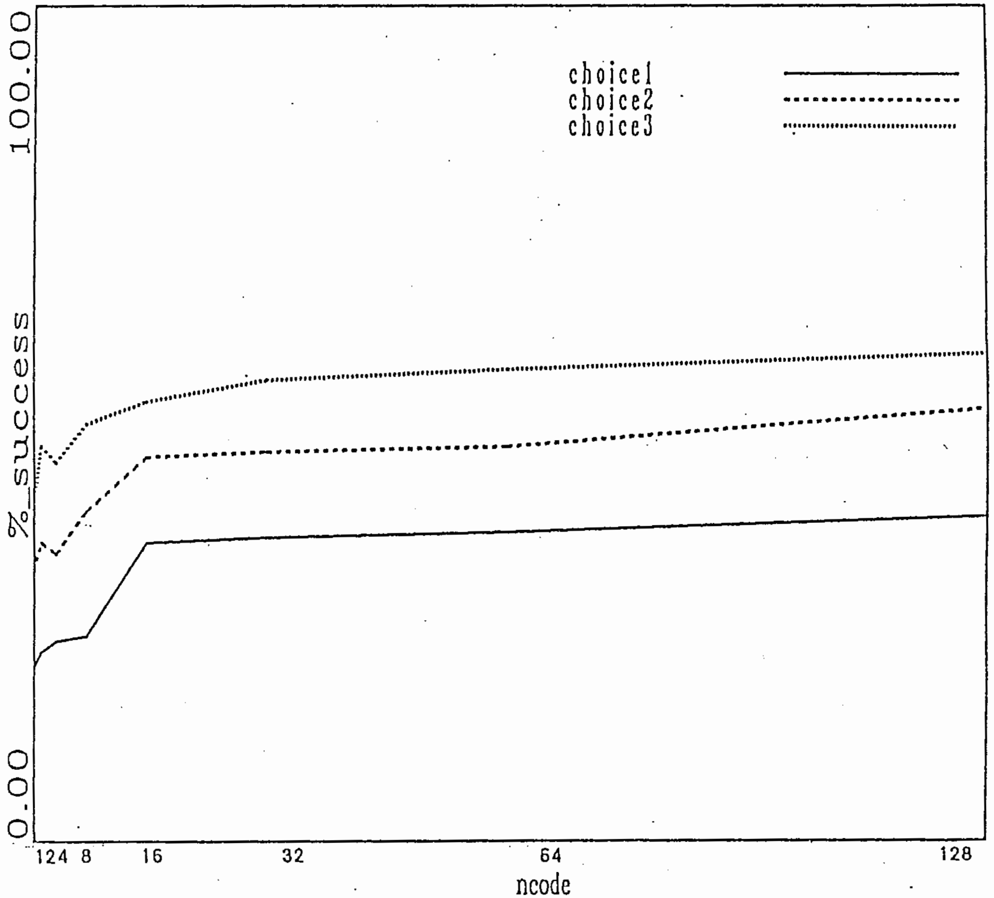


Figure 4: The influence of number of VQ codebook

parameter	value
number of codes	16
distance	WLR
amount of test data	1
frame shift for test	512

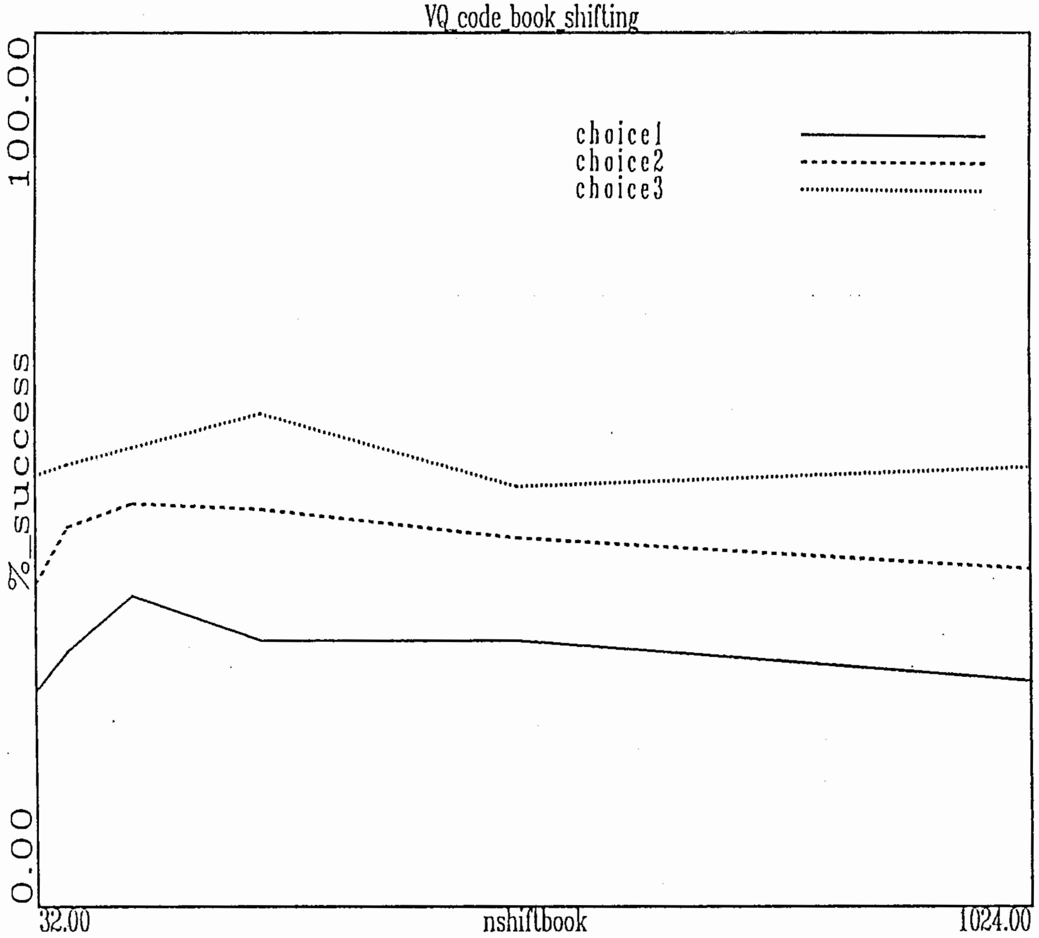


Figure 5: The influence of shifting in training

parameter	value
number of codes	16
distance	WLR
amount of test data	1
frame shift for training	128

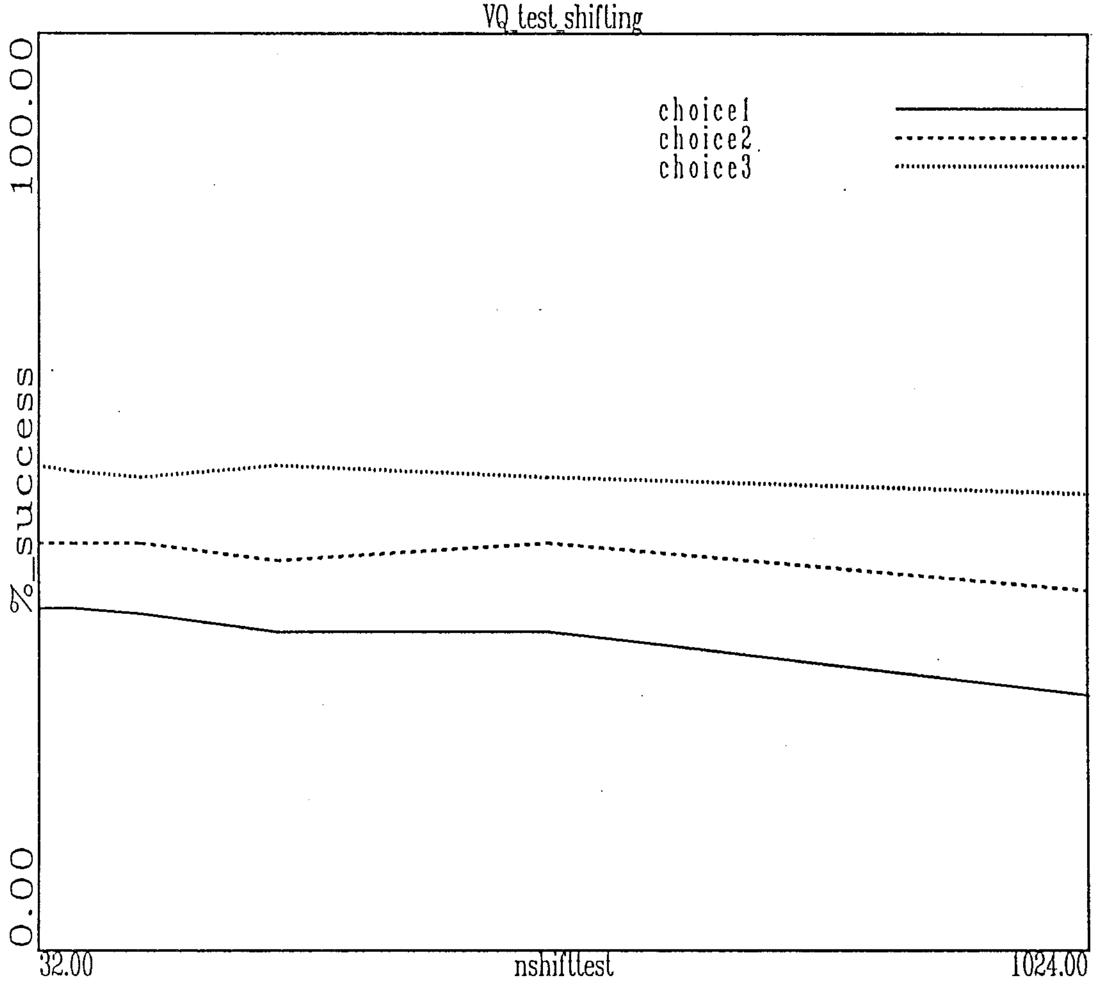


Figure 6: The influence of shifting in testing

parameter	value
number of codes	64
distance	CEP+ΔCEP
amount of test data	8
frame shift for training	128
frame shift for test	512

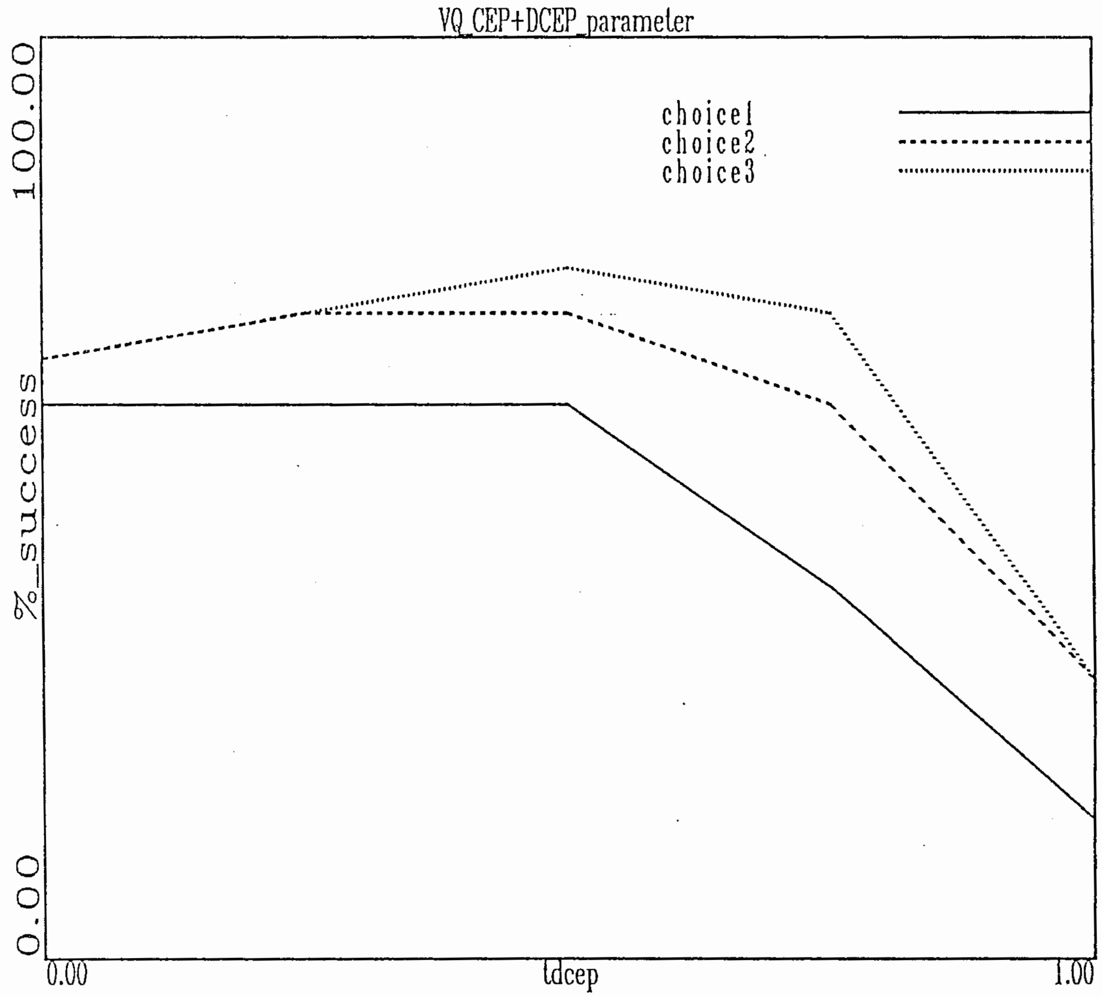


Figure 7: CEP & ΔCEP

parameter	value
number of codes	16
frame shift for training	128
frame shift for test	64
CEP+ΔCEP parameter	0.5

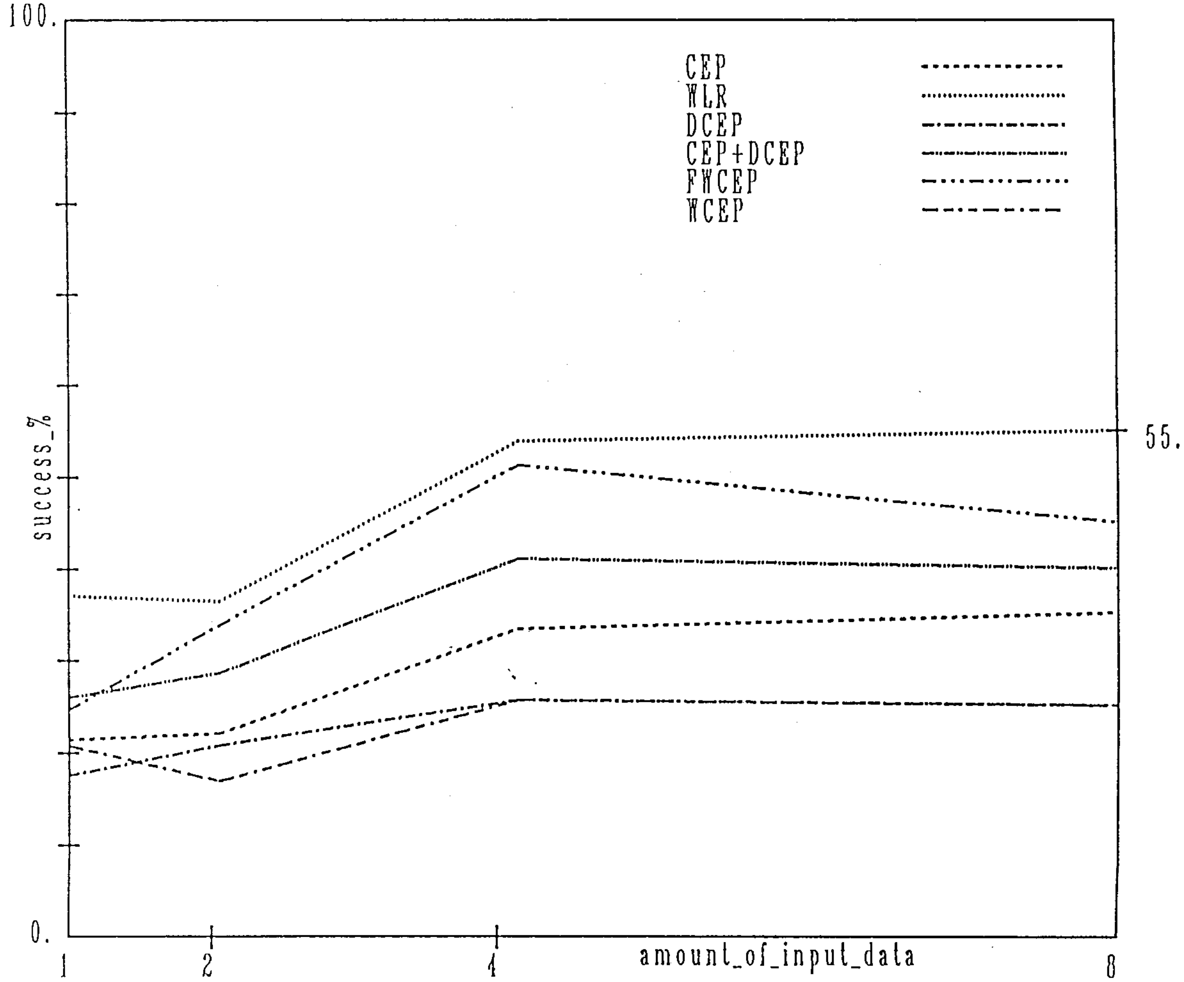


Figure 8: Recognition accuracy for different distances

parameter	value
number of codes	64
frame shift for training	128
frame shift for test	64

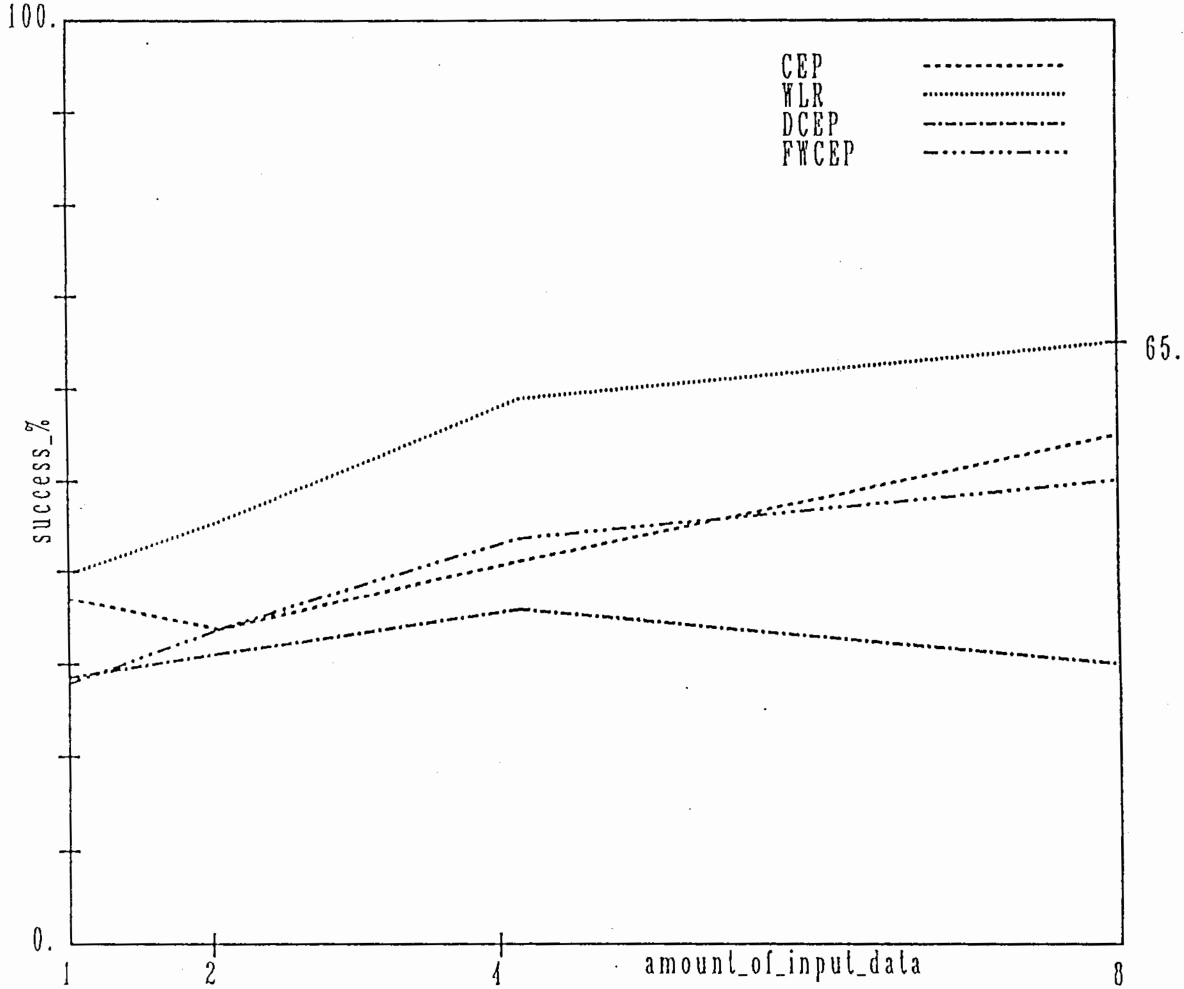


Figure 9: Recognition accuracy for different distances

CEP

LANGUAGE	CH 1	CH 2	CH 3	CH 5	CH 10
AM	100.00	100.00	100.00	100.00	100.00
AR	0.00	0.00	0.00	0.00	50.00
CH	0.00	0.00	0.00	0.00	50.00
DA	50.00	50.00	100.00	100.00	100.00
DU	50.00	50.00	50.00	100.00	100.00
EN	50.00	100.00	100.00	100.00	100.00
FI	0.00	50.00	100.00	100.00	100.00
FR	50.00	100.00	100.00	100.00	100.00
GE	0.00	0.00	0.00	50.00	50.00
GR	0.00	0.00	0.00	100.00	100.00
HI	0.00	0.00	0.00	50.00	50.00
HU	0.00	50.00	100.00	100.00	100.00
IT	0.00	0.00	0.00	50.00	100.00
JA	100.00	100.00	100.00	100.00	100.00
NO	50.00	50.00	100.00	100.00	100.00
PL	100.00	100.00	100.00	100.00	100.00
PR	100.00	100.00	100.00	100.00	100.00
RU	0.00	0.00	0.00	0.00	0.00
SP	0.00	50.00	50.00	50.00	100.00
SW	50.00	50.00	50.00	100.00	100.00
TOTAL	33.33	46.15	56.41	74.36	84.62

WLR

LANGUAGE	CH 1	CH 2	CH 3	CH 5	CH 10
AM	100.00	100.00	100.00	100.00	100.00
AR	0.00	0.00	0.00	50.00	50.00
CH	100.00	100.00	100.00	100.00	100.00
DA	100.00	100.00	100.00	100.00	100.00
DU	100.00	100.00	100.00	100.00	100.00
EN	50.00	50.00	50.00	100.00	100.00
FI	50.00	50.00	100.00	100.00	100.00
FR	0.00	100.00	100.00	100.00	100.00
GE	50.00	50.00	50.00	50.00	50.00
GR	0.00	0.00	50.00	100.00	100.00
HI	0.00	0.00	0.00	0.00	0.00
HU	50.00	50.00	50.00	100.00	100.00
IT	50.00	50.00	100.00	100.00	100.00
JA	100.00	100.00	100.00	100.00	100.00
NO	100.00	100.00	100.00	100.00	100.00
PL	100.00	100.00	100.00	100.00	100.00
PR	100.00	100.00	100.00	100.00	100.00
RU	0.00	0.00	0.00	50.00	50.00
SP	0.00	50.00	50.00	100.00	100.00
SW	50.00	50.00	50.00	100.00	100.00
TOTAL	53.85	61.54	69.23	87.18	87.18

Figure 10: Recognition accuracy with several distances

CEP+ΔCEP

LANGUAGE	CH 1	CH 2	CH 3	CH 5	CH 10
AM	100.00	100.00	100.00	100.00	100.00
AR	0.00	0.00	0.00	50.00	50.00
CH	0.00	0.00	0.00	0.00	100.00
DA	50.00	50.00	50.00	100.00	100.00
DU	0.00	50.00	50.00	50.00	100.00
EN	100.00	100.00	100.00	100.00	100.00
FI	50.00	50.00	50.00	100.00	100.00
FR	100.00	100.00	100.00	100.00	100.00
GE	0.00	0.00	0.00	50.00	50.00
GR	0.00	0.00	0.00	100.00	100.00
HI	0.00	0.00	50.00	50.00	100.00
HU	0.00	50.00	50.00	50.00	100.00
IT	0.00	0.00	0.00	100.00	100.00
JA	100.00	100.00	100.00	100.00	100.00
NO	50.00	50.00	100.00	100.00	100.00
PL	100.00	100.00	100.00	100.00	100.00
PR	100.00	100.00	100.00	100.00	100.00
RU	0.00	0.00	0.00	0.00	0.00
SP	50.00	50.00	50.00	100.00	100.00
SW	50.00	50.00	100.00	100.00	100.00
TOTAL	41.03	46.15	53.85	76.92	89.74

FWCEP

LANGUAGE	CH 1	CH 2	CH 3	CH 5	CH 10
AM	100.00	100.00	100.00	100.00	100.00
AR	0.00	0.00	0.00	50.00	50.00
CH	0.00	0.00	50.00	50.00	50.00
DA	0.00	0.00	0.00	50.00	100.00
DU	50.00	100.00	100.00	100.00	100.00
EN	100.00	100.00	100.00	100.00	100.00
FI	0.00	0.00	0.00	50.00	100.00
FR	100.00	100.00	100.00	100.00	100.00
GE	50.00	50.00	50.00	50.00	50.00
GR	50.00	100.00	100.00	100.00	100.00
HI	0.00	0.00	100.00	100.00	100.00
HU	50.00	50.00	50.00	50.00	100.00
IT	100.00	100.00	100.00	100.00	100.00
JA	100.00	100.00	100.00	100.00	100.00
NO	100.00	100.00	100.00	100.00	100.00
PL	100.00	100.00	100.00	100.00	100.00
PR	0.00	50.00	50.00	100.00	100.00
RU	50.00	50.00	50.00	50.00	50.00
SP	0.00	0.00	0.00	100.00	100.00
SW	100.00	100.00	100.00	100.00	100.00
TOTAL	51.28	58.97	66.67	82.05	89.74

Figure 10-b: Recognition accuracy with several distances

parameter	value
number of codes	16
distance	CEP
amount of test data	4
frame shift for training	128
frame shift for test	64

CEP

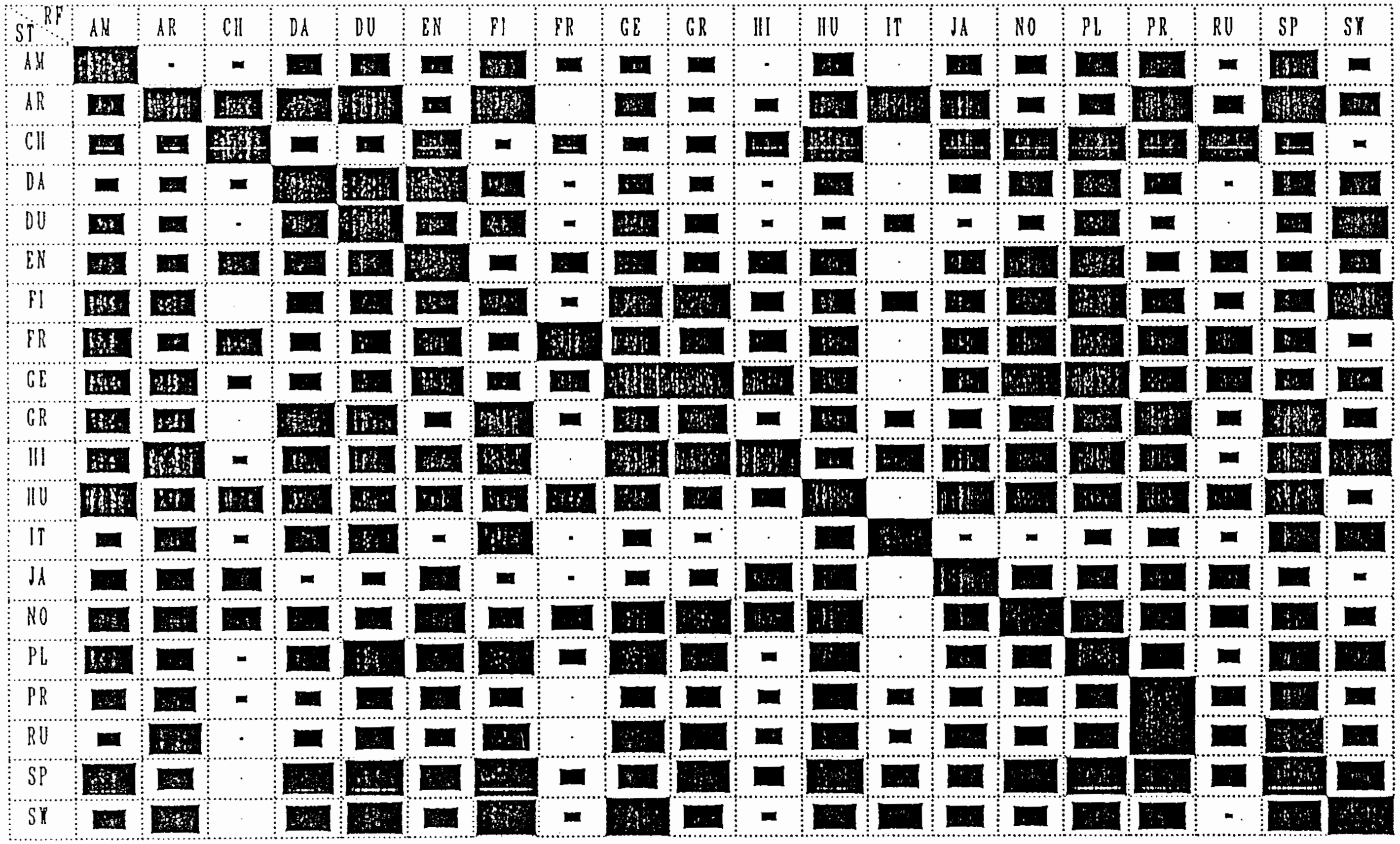


Figure 12: Confusion matrix for standard VQ algorithm (CEP)

parameter	value
number of codes	16
distance	WLR
amount of test data	4
frame shift for training	128
frame shift for test	64

WLR

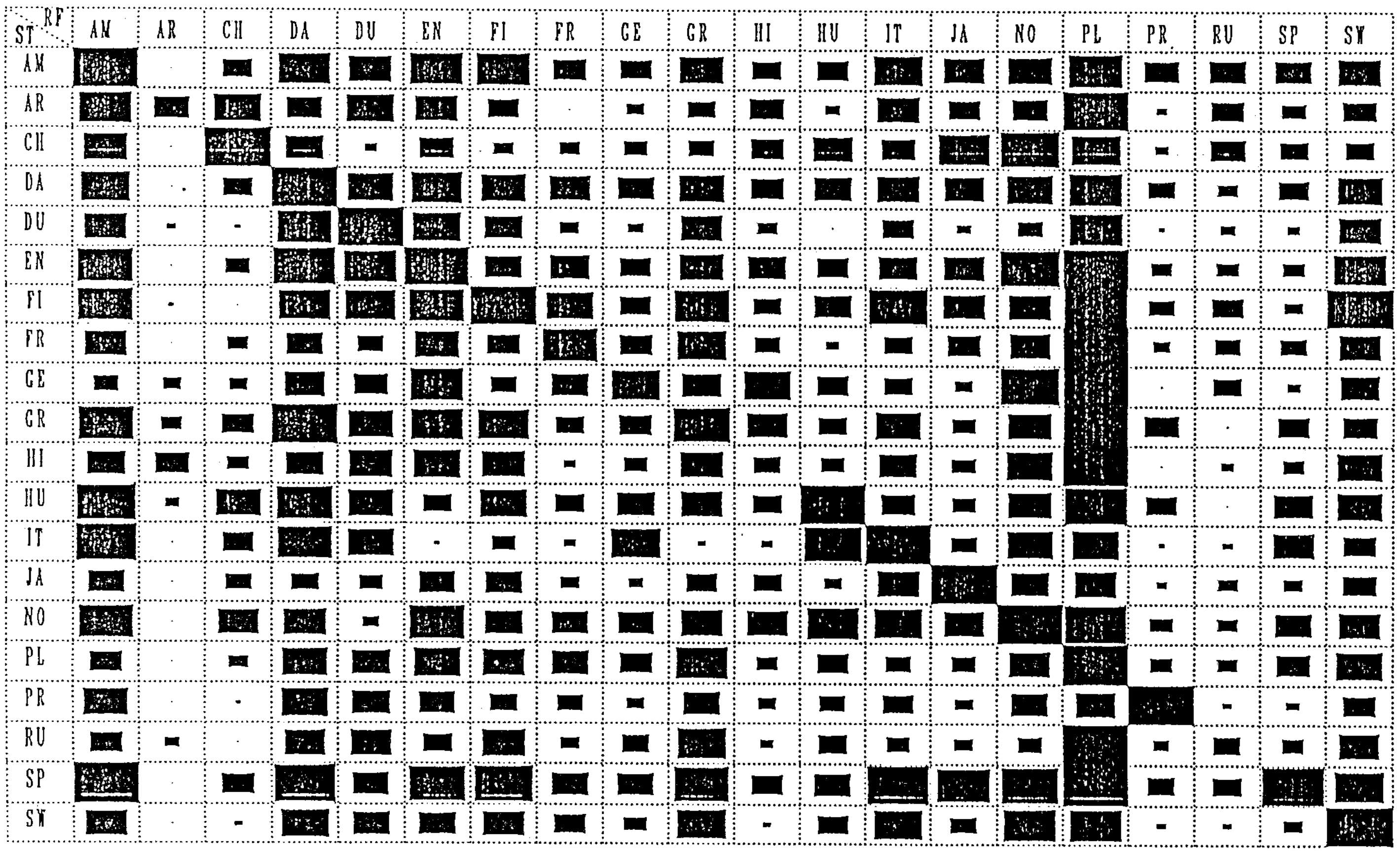


Figure 11: Confusion matrix for standard VQ algorithm (WLR)

parameter	value
distance	WLR
amount of test data	1
frame shift for training	1092
frame shift for test	512

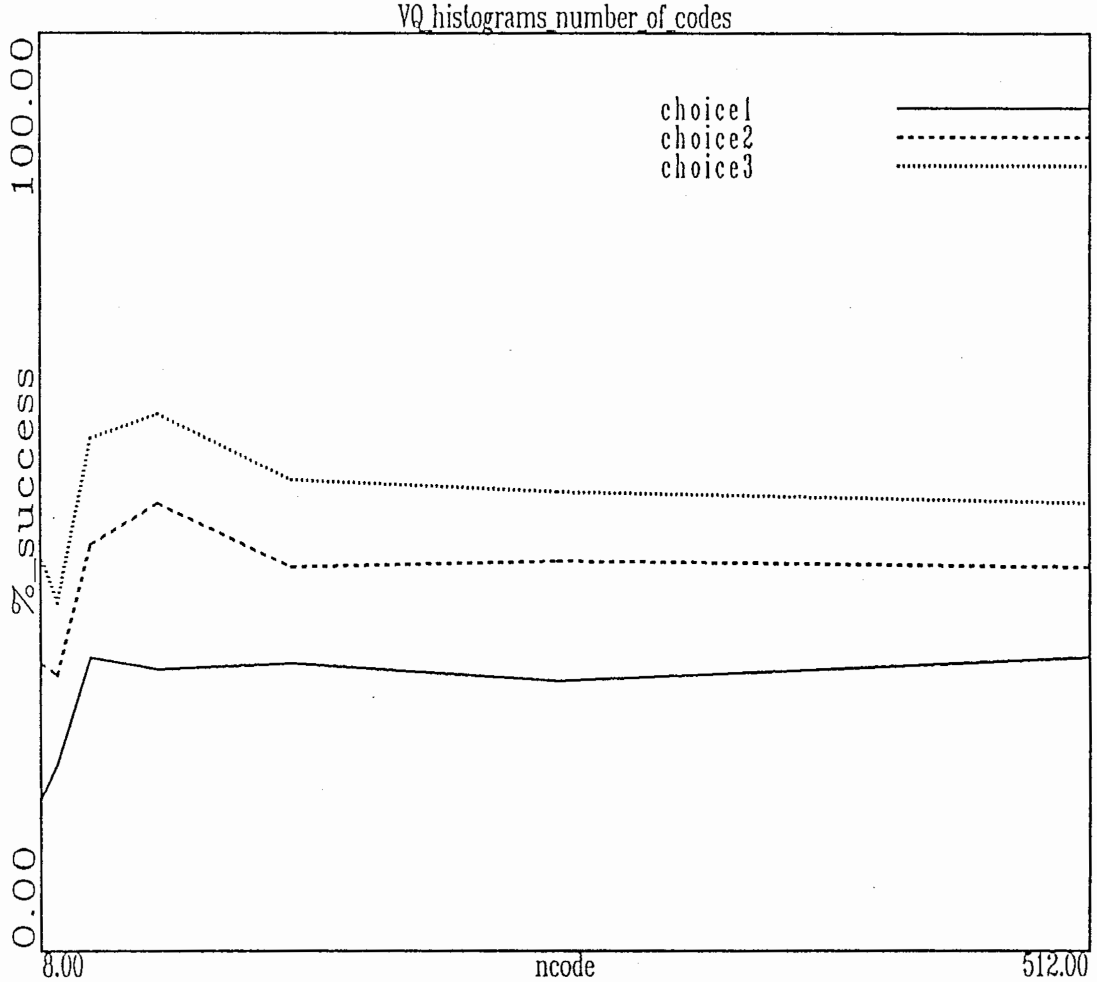


Figure 13: Number of VQ codebook in VQ histogram algorithm

parameter	value
number of codes	32
distance	WLR
amount of test data	1
frame shift for test	512

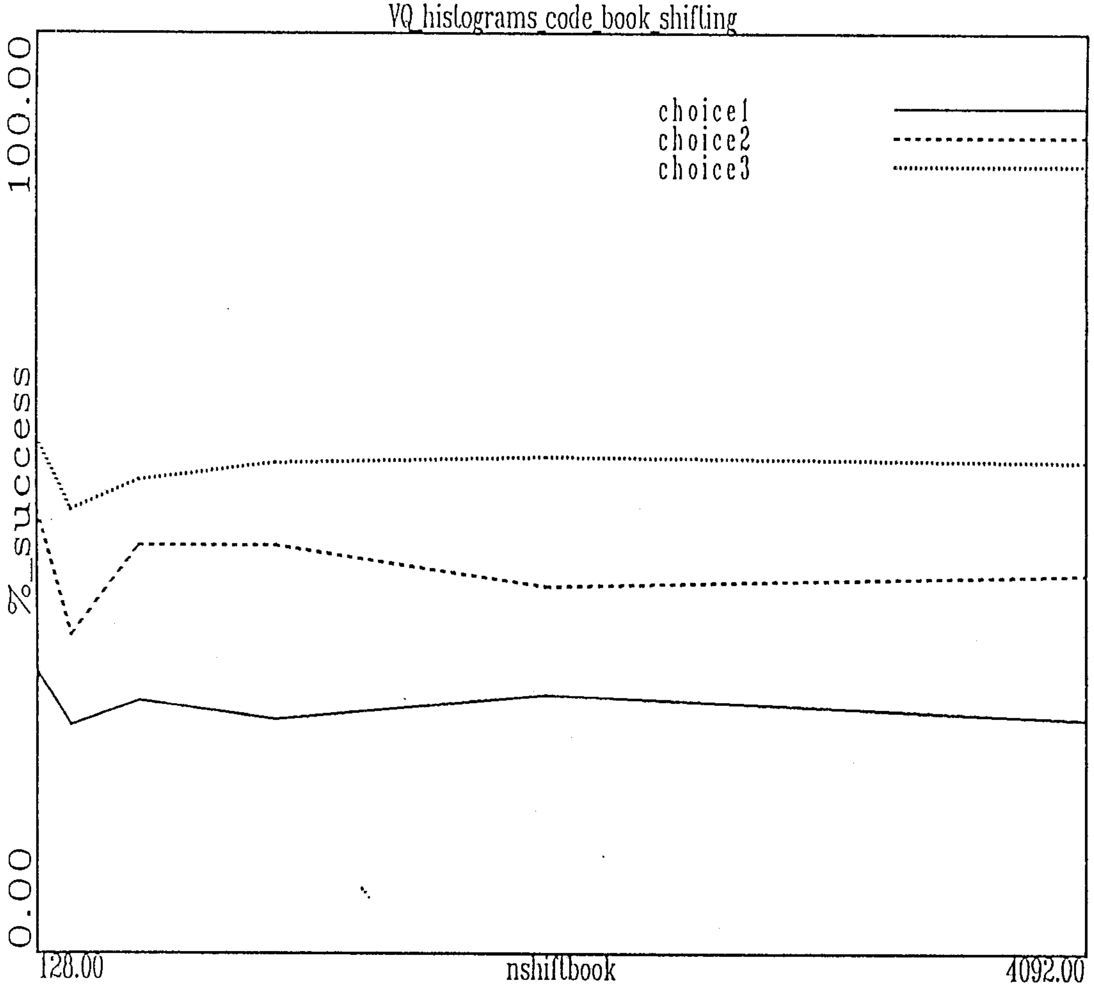


Figure 14: The influence of shifting in training (VQ histogram algorithm)

parameter	value
number of codes	32
distance	WLR
amount of test data	1
frame shift for training	128

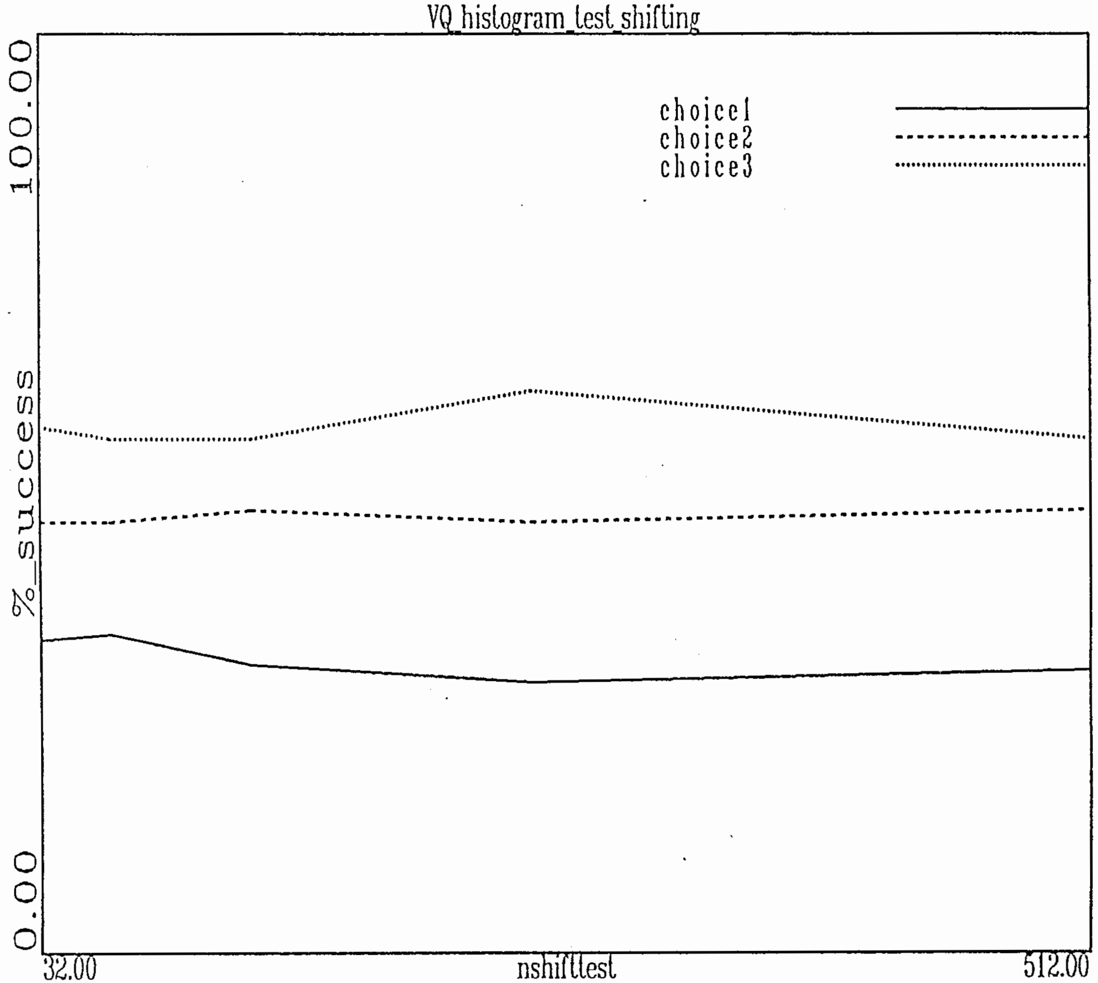


Figure 15: The influence of shifting in testing (VQ histogram algorithm)

parameter	value
number of codes	32
frame shift for training	128
frame shift for test	64

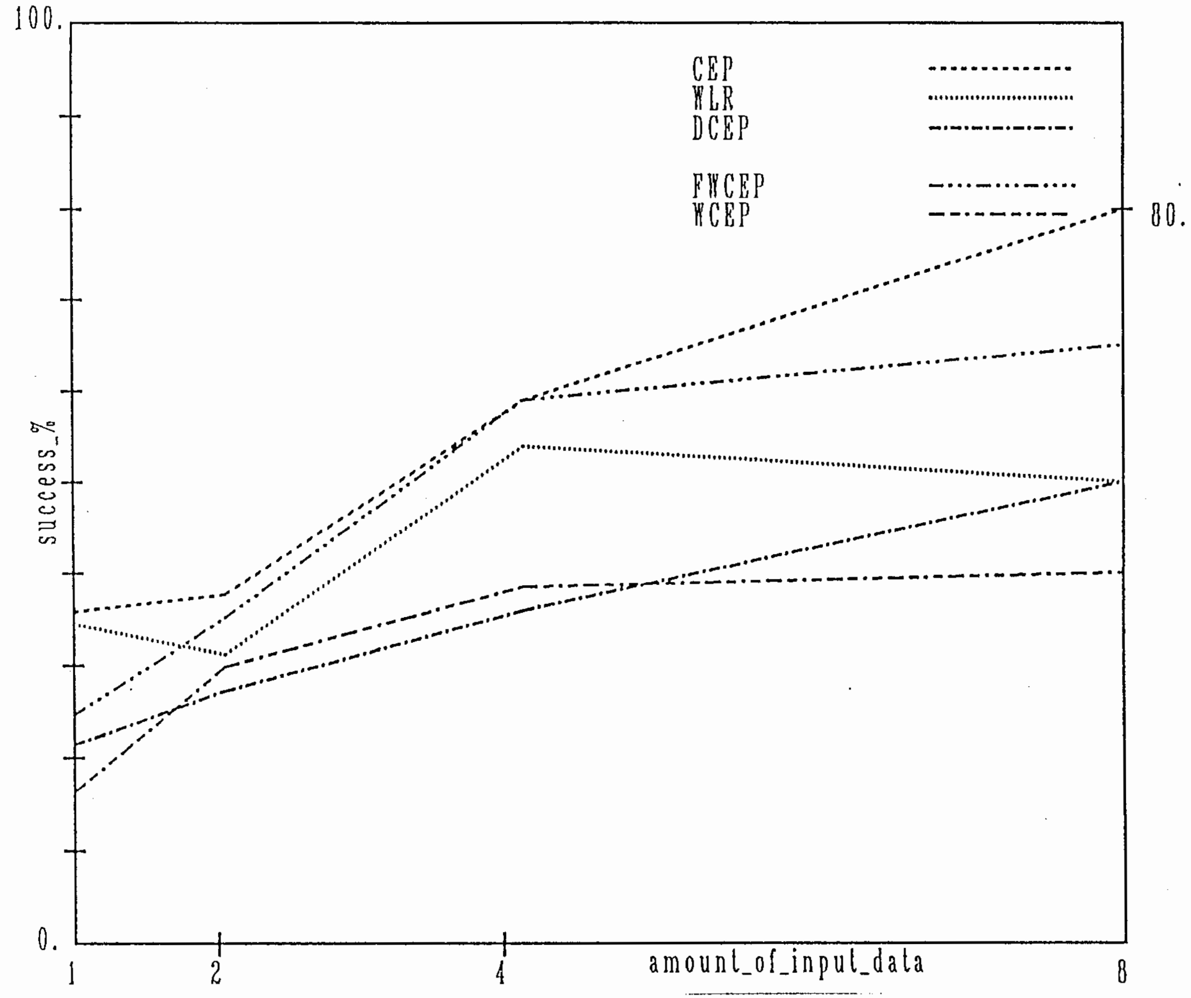


Figure 16: VQ histogram algorithm using different distances

parameter	value
number of codes	64
frame shift for training	128
frame shift for test	64

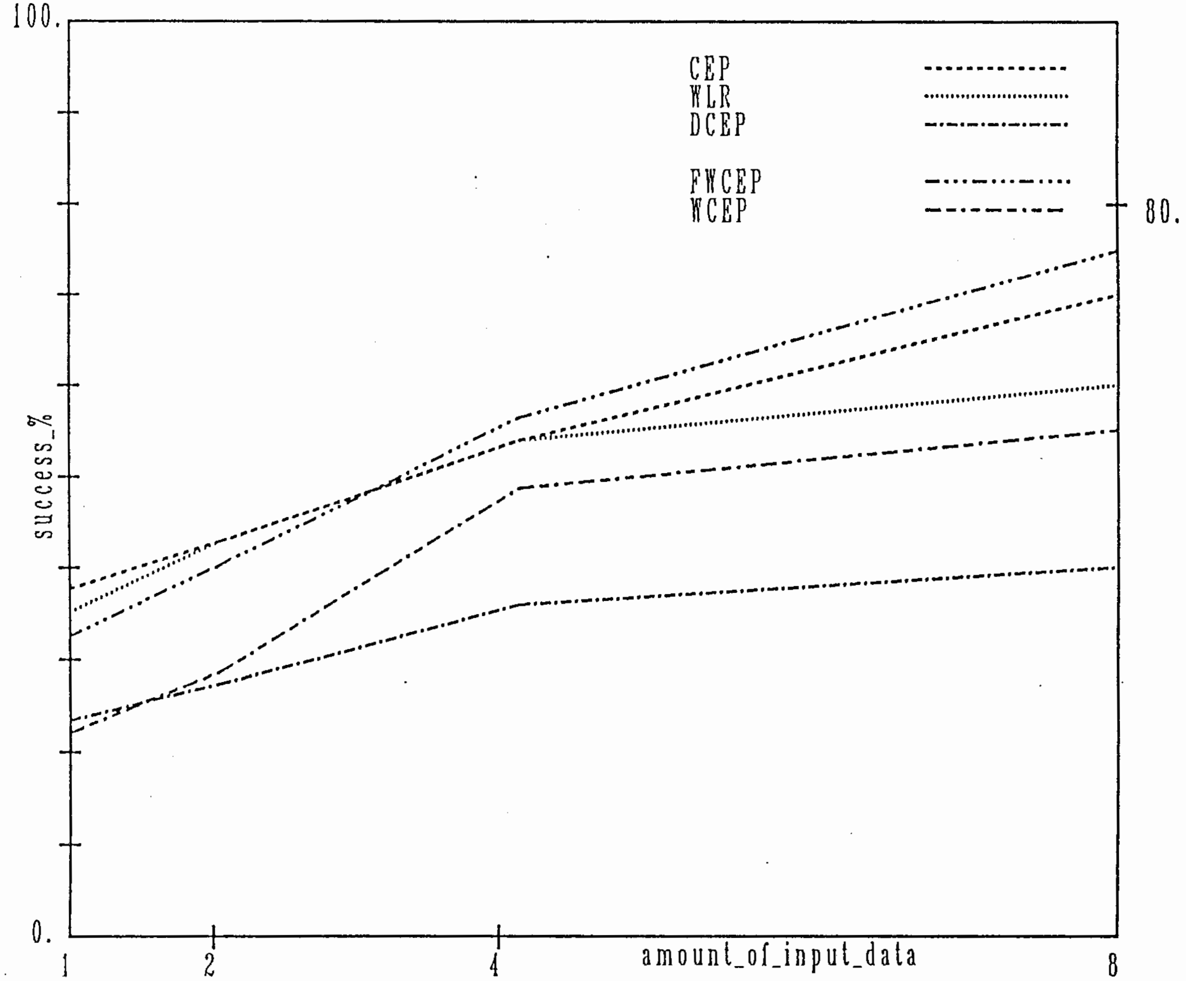
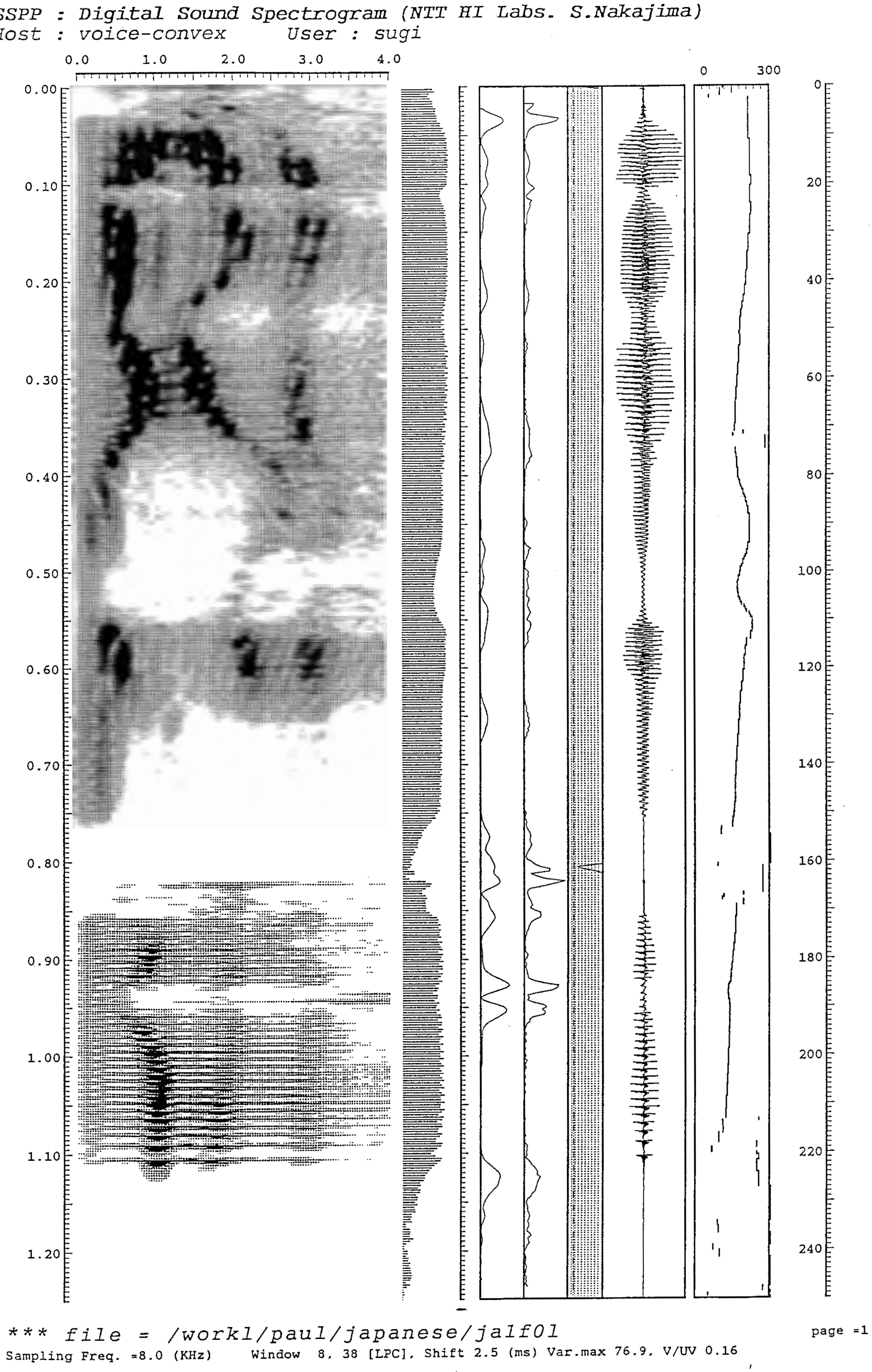


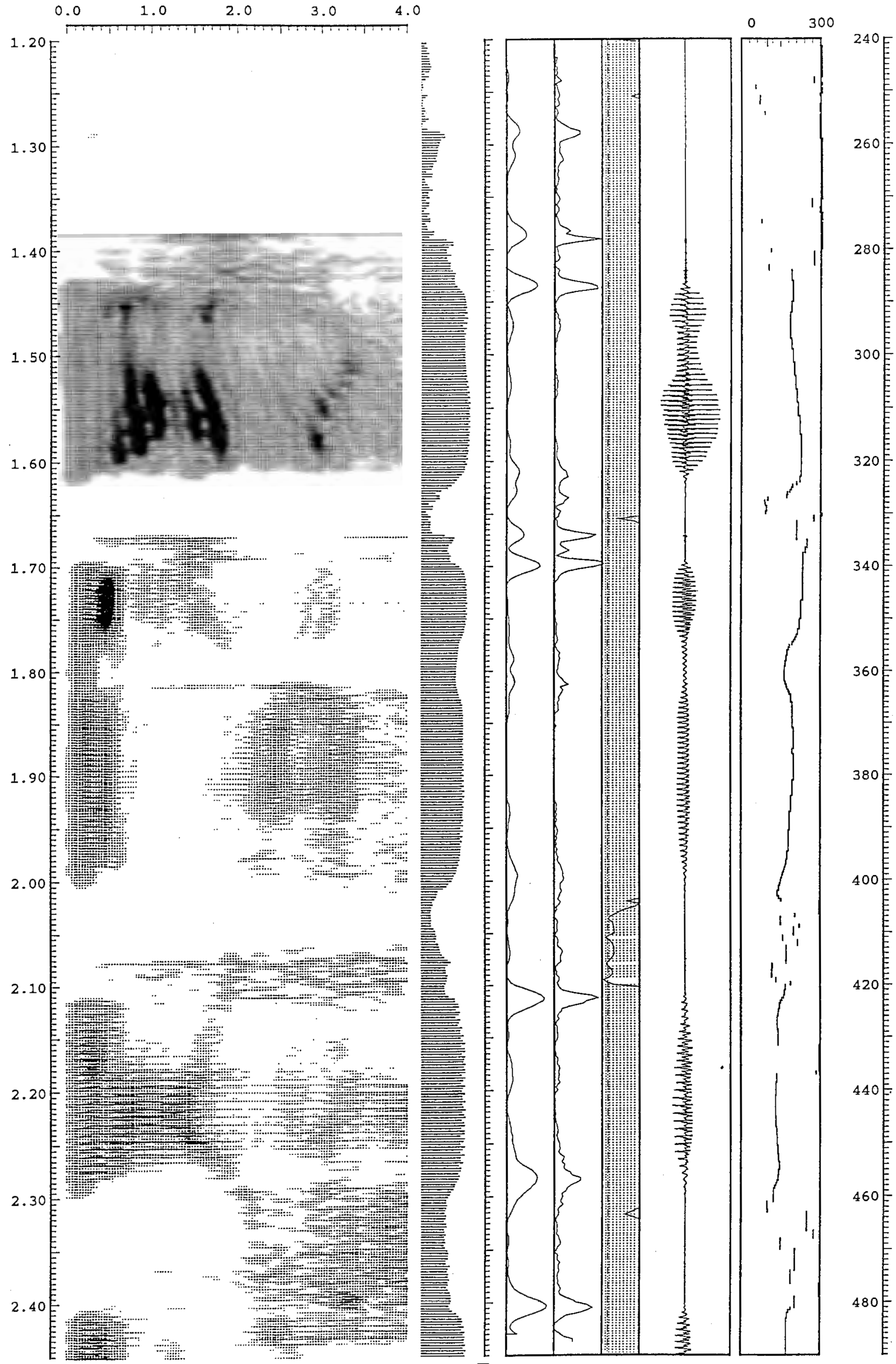
Figure 17: VQ histogram algorithm using different distances

J.2 Sonagram for several sentences

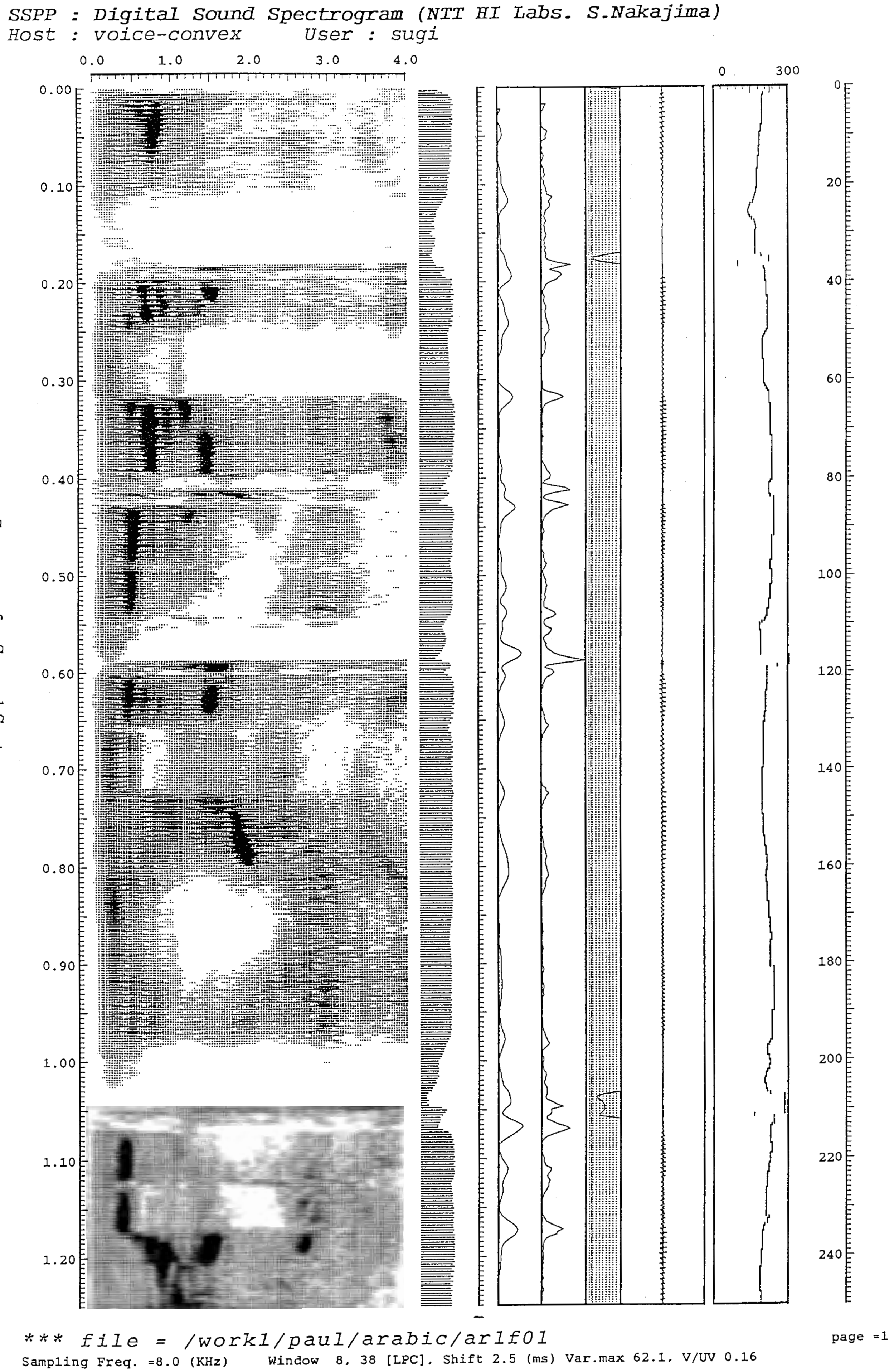


Sonagram for Several Sentences

Host : voice-convex   User : sugi

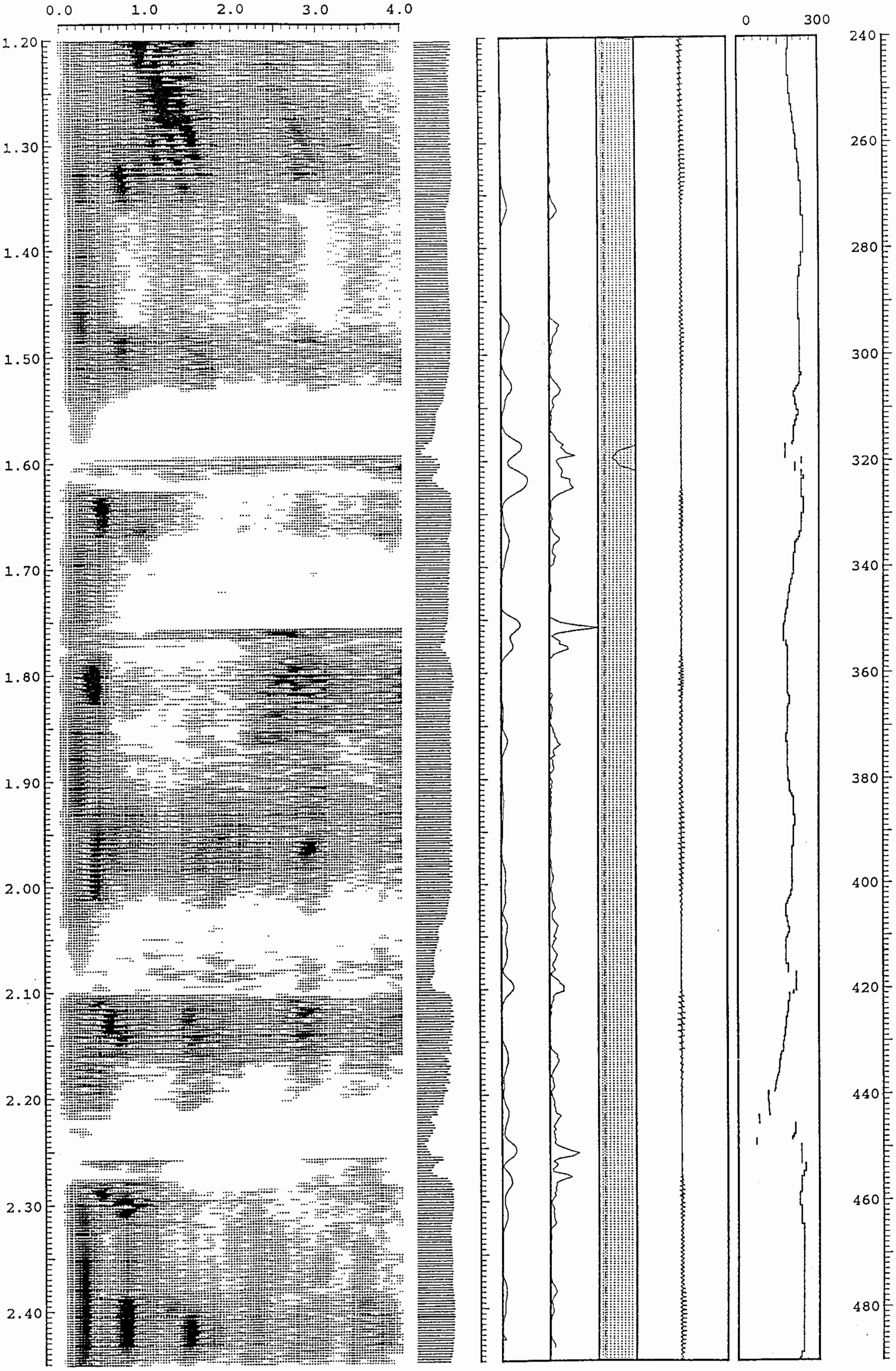


*** file = /work1/paul/japanese/ja1f01
Sampling Freq. =8.0 (KHz)   Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 76.9, V/UV 0.16



Sonagram for Several Sentences

Host : voice-convex User : sugi

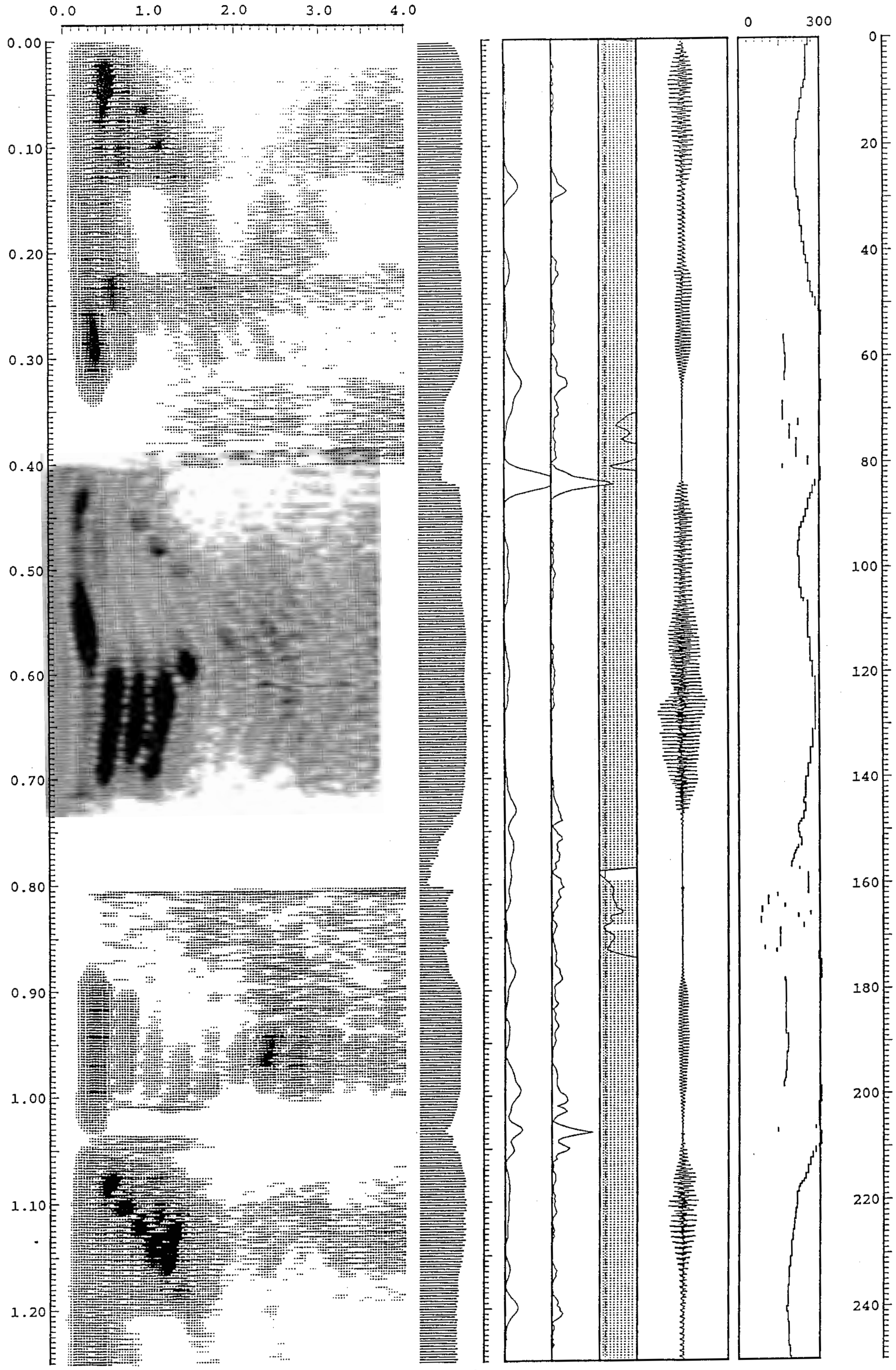


*** file = /work1/paul/arabic/ar1f01

page =2

Sampling Freq. =8.0 (KHz) Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 62.1, V/UV 0.16

SSPP : Digital Sound Spectrogram (NTT HI Labs. S.Nakajima)
Host : voice-convex User : sugi



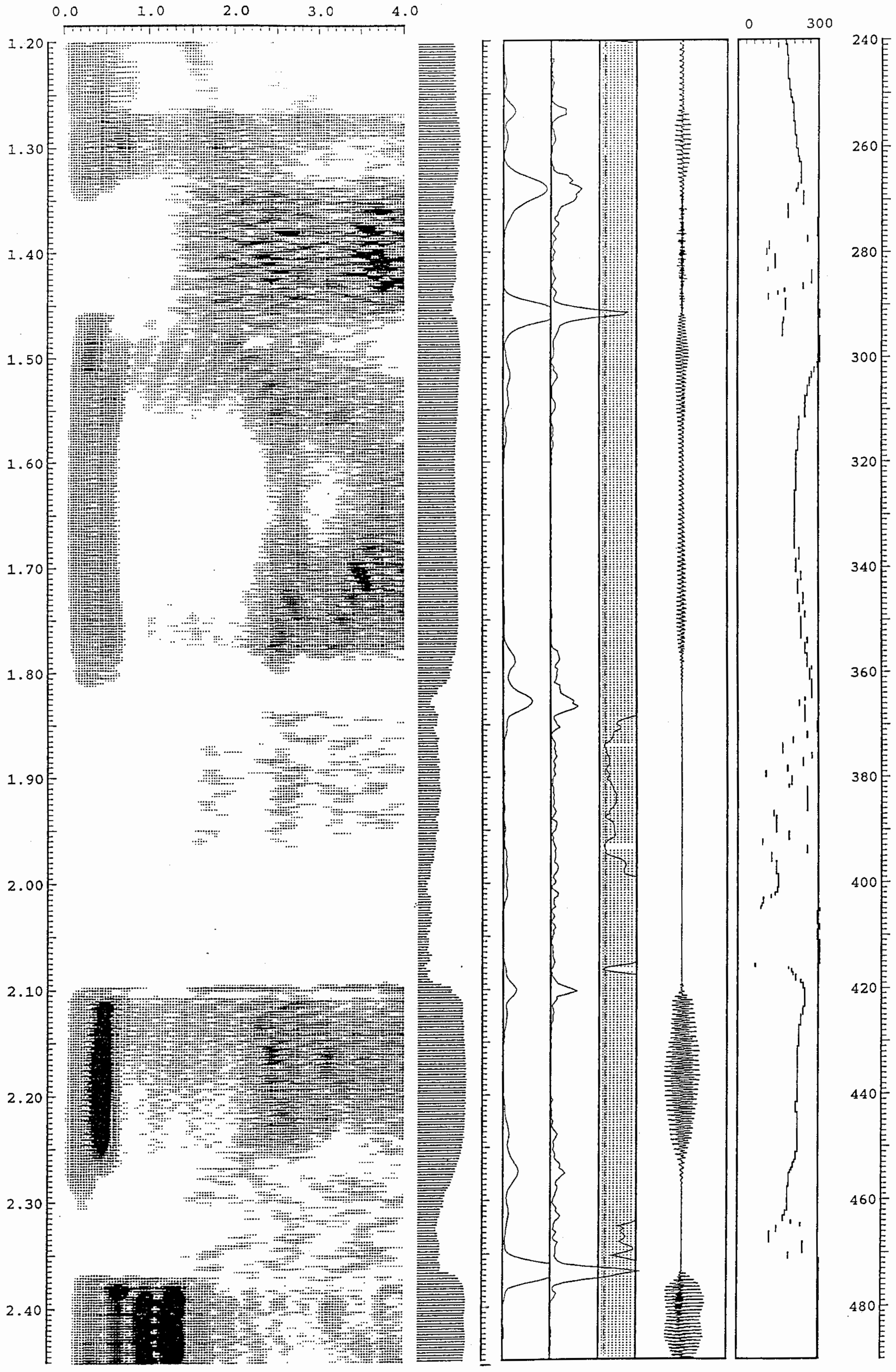
*** file = /work1/paul/chinese/ch1f02

page =1

Sampling Freq. =8.0 (KHz) Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 76.7, V/UV 0.16

Sonagram for Several Sentences

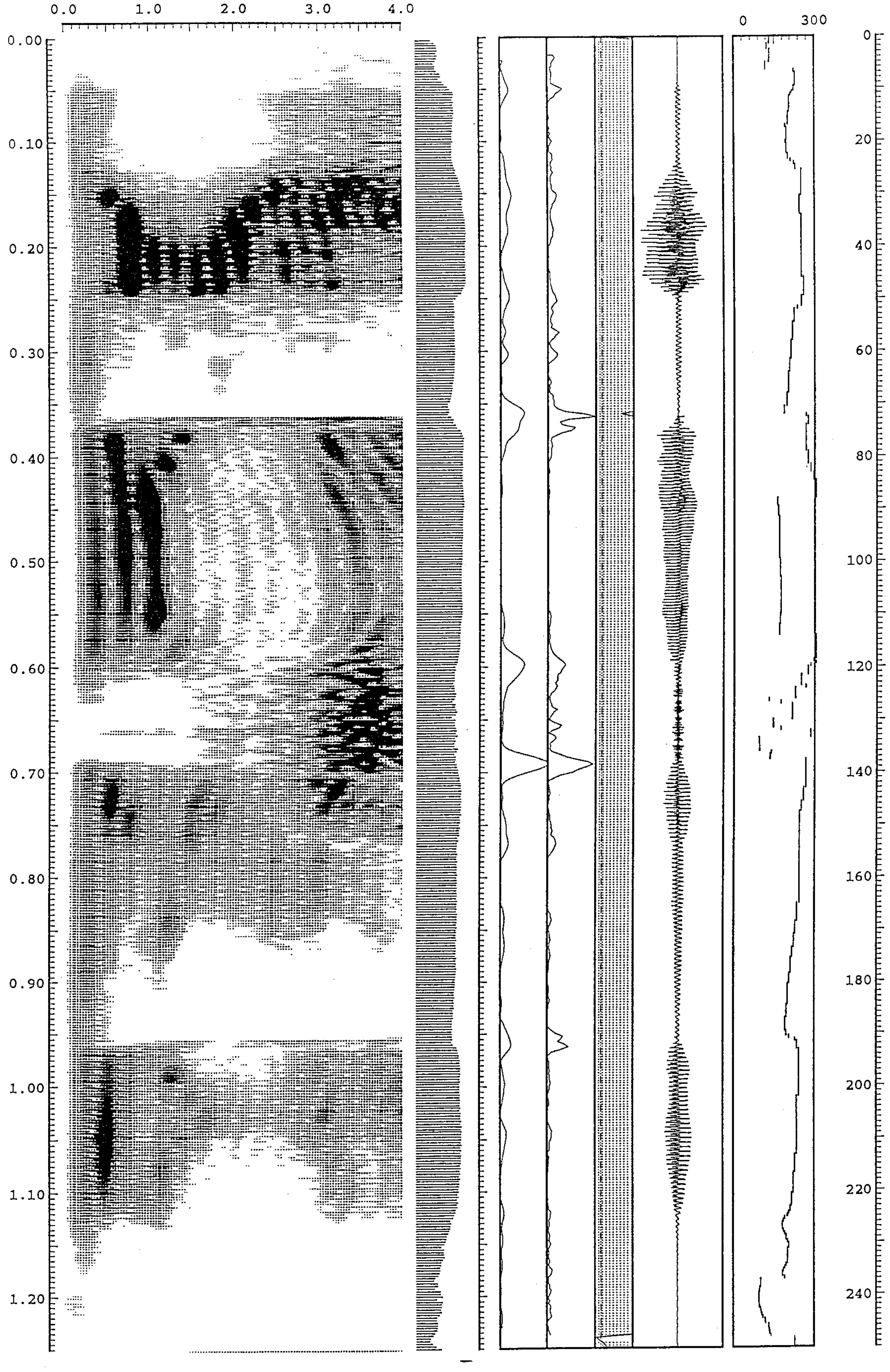
Host : voice-convex   User : sugi



*** file = /work1/paul/chinese/ch1f02   page =2

Sampling Freq. =8.0 (KHz)   Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 76.7, V/UV 0.16

SSPP : Digital Sound Spectrogram (NTT HI Labs. S.Nakajima)
Host : voice-convex User : sugi



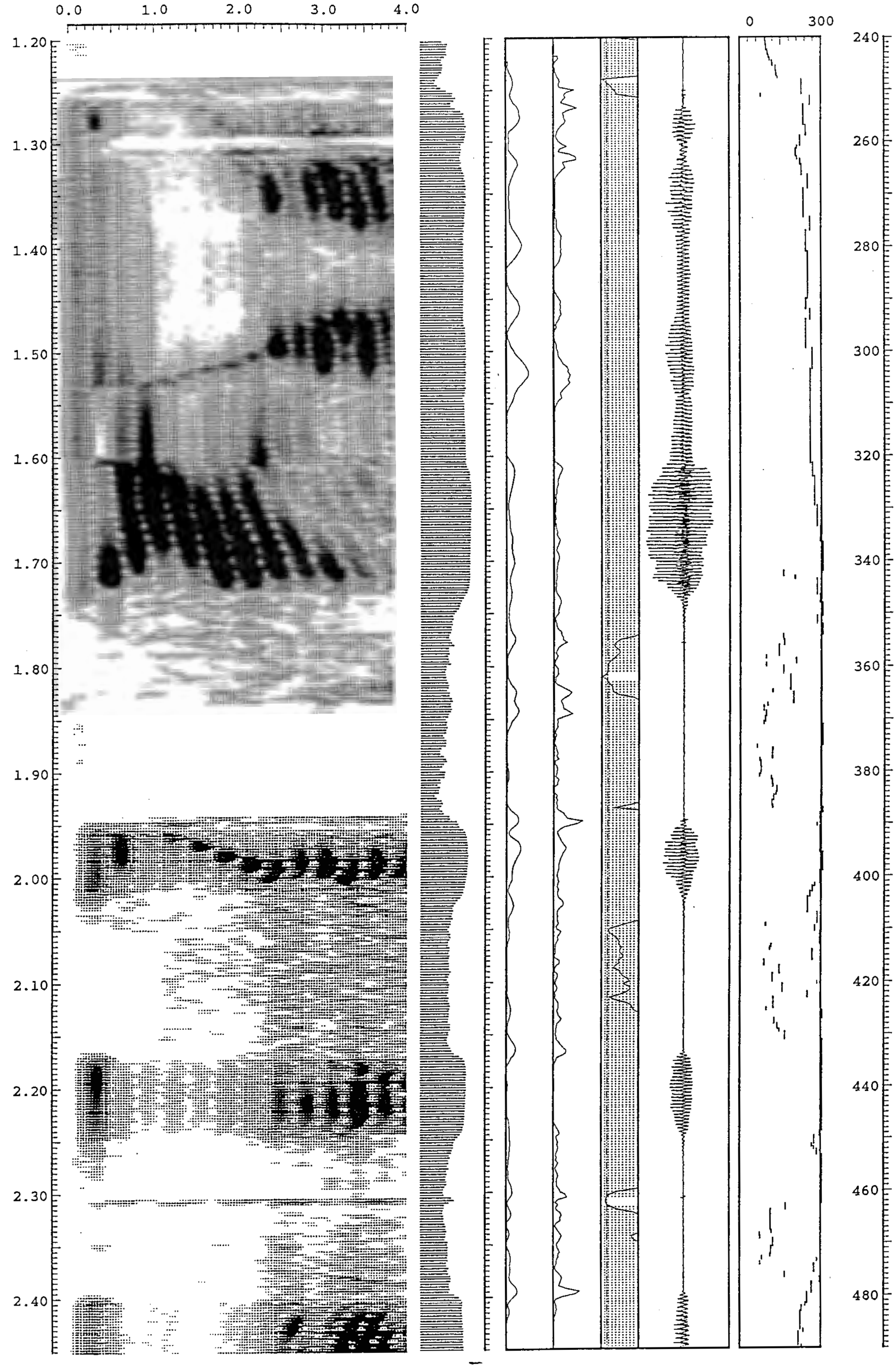
Sonagram for Several Sentences

**** file = /work1/paul/russian/ru1f09*

page =1

Sampling Freq. =8.0 (KHz) Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 76.7, V/UV 0.16

Host : voice-convex User : sugi

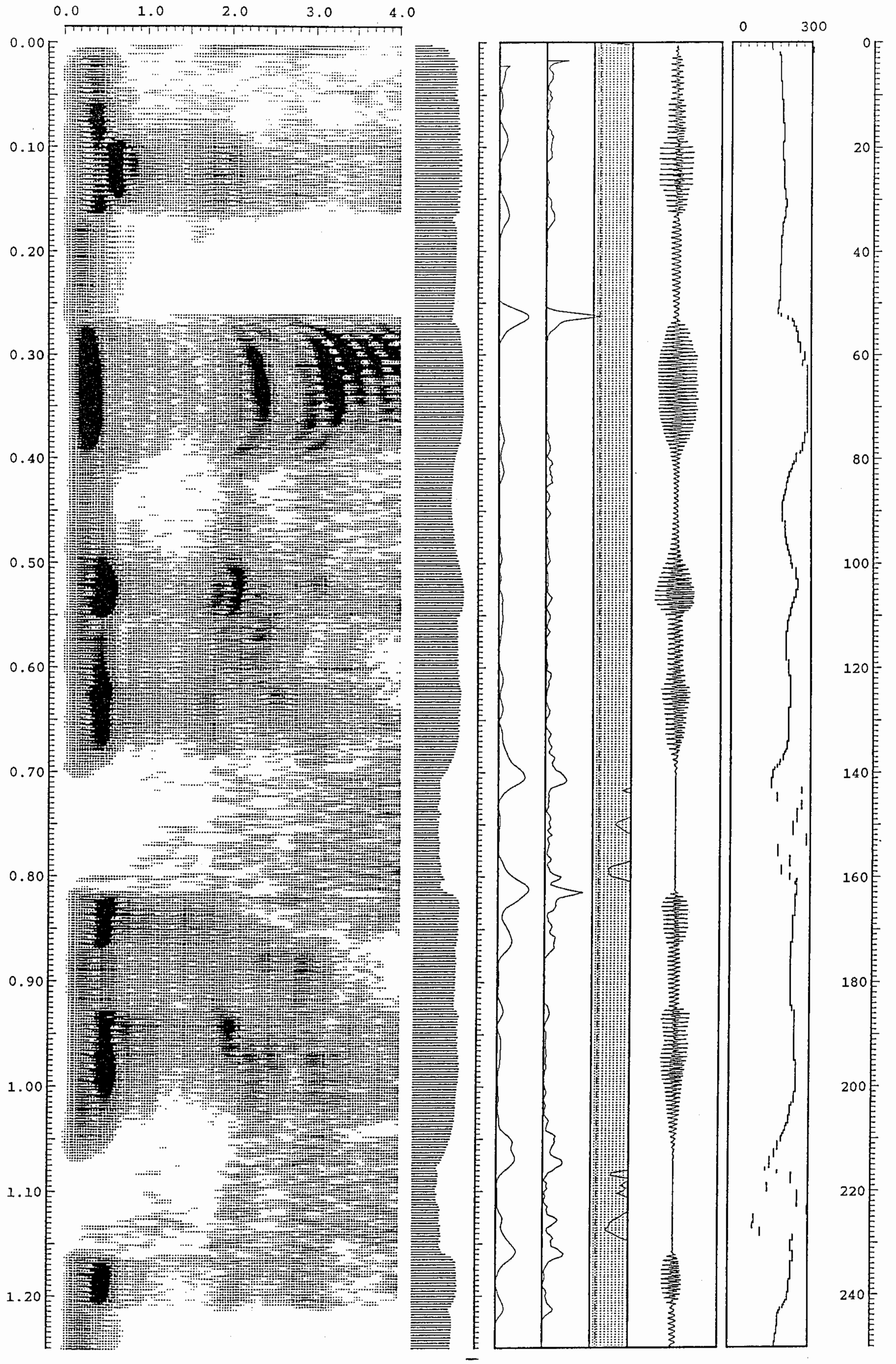


*** file = /work1/paul/russian/ru1f09

page =2

Sampling Freq. =8.0 (KHz) Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 76.7, V/UV 0.16

SSPP : Digital Sound Spectrogram (NTT HI Labs. S.Nakajima)
Host : voice-convex   User : sugi

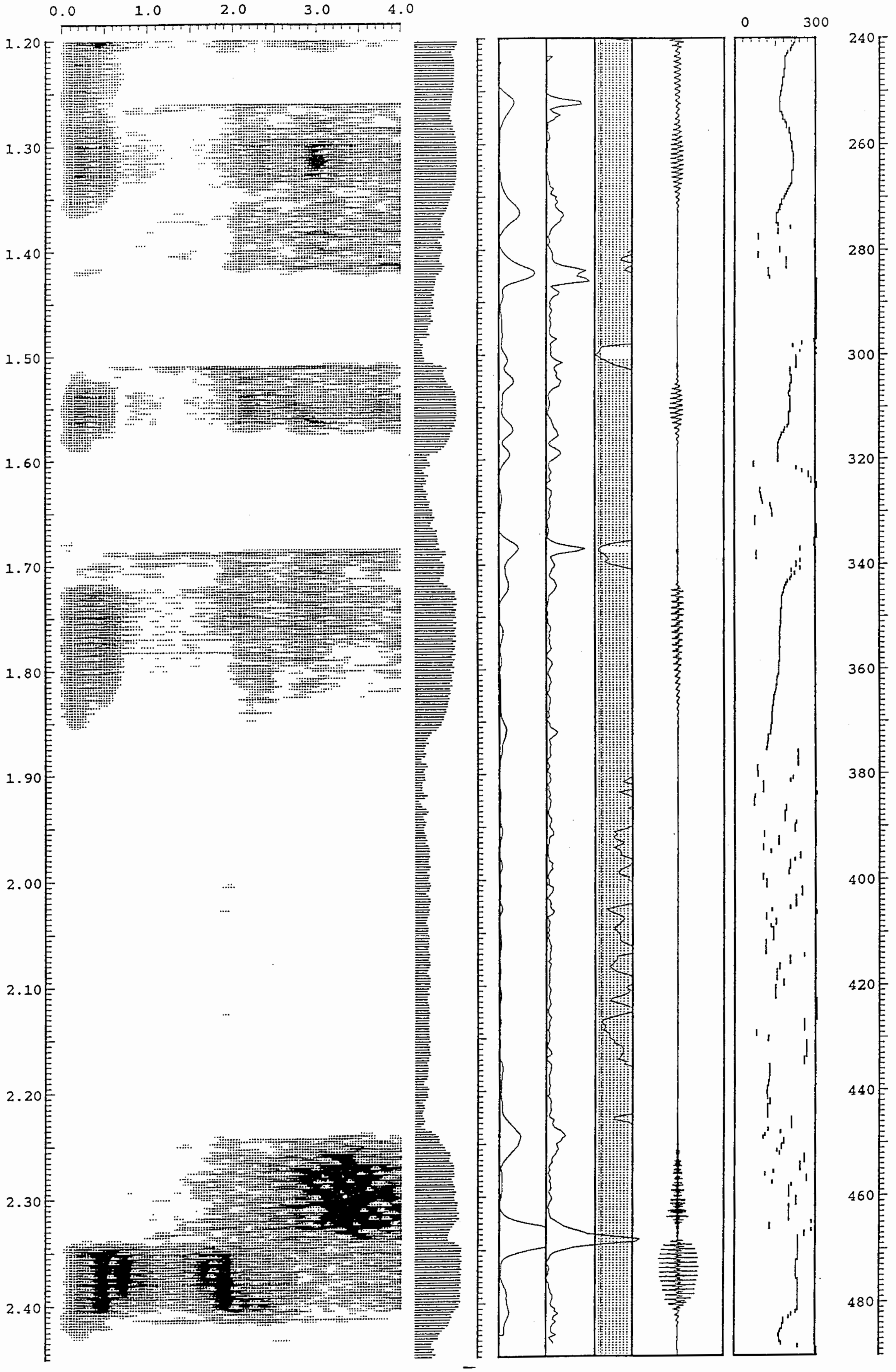


Sonagram for Several Sentences

*** file = /work1/paul/french/fr1f01   page =1

Sampling Freq. =8.0 (KHz)   Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 74.6, V/UV 0.16

Host : voice-convex   User : sugi



*** file = /work1/paul/french/fr1f01

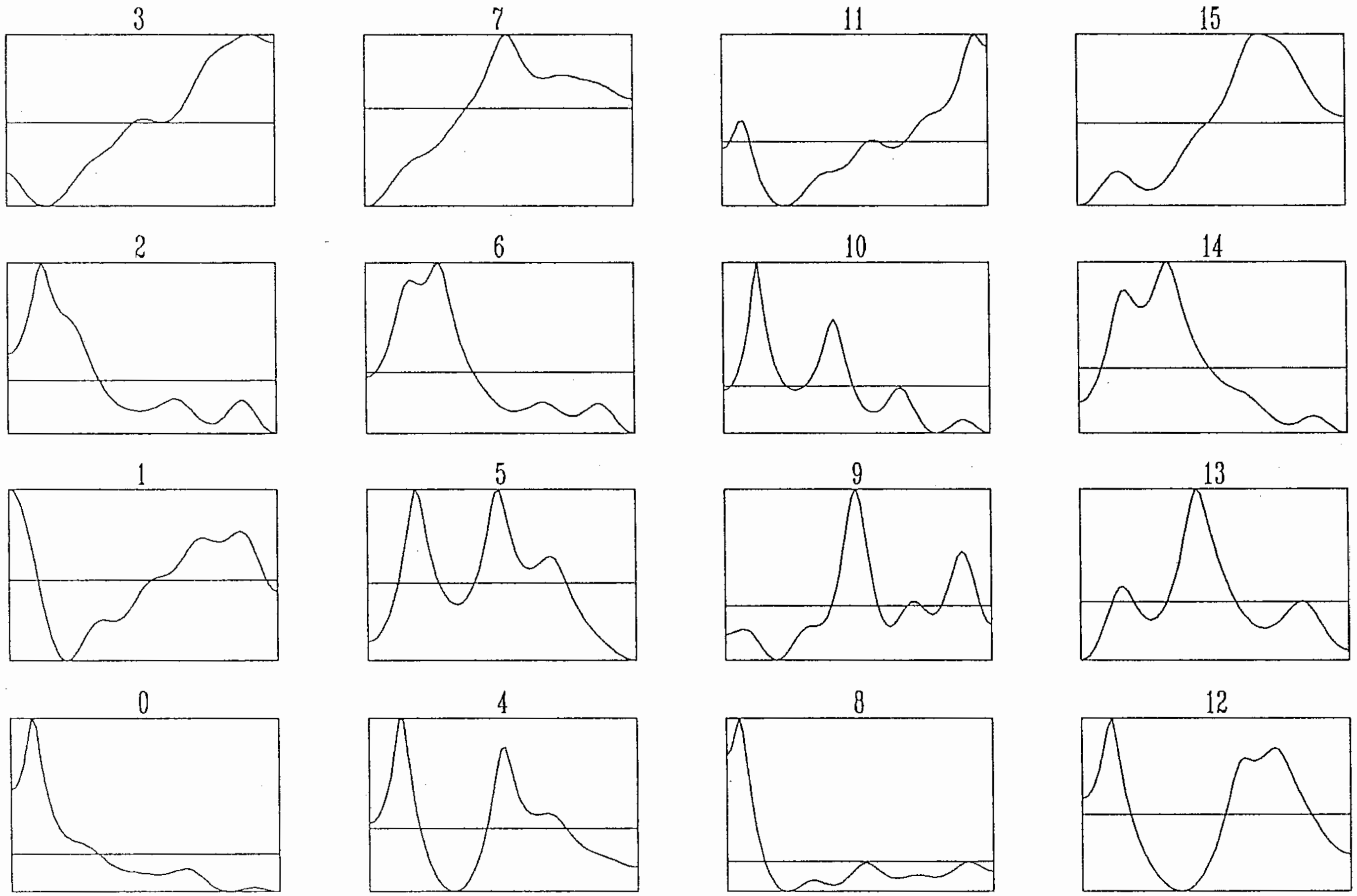
page =2

Sampling Freq. =8.0 (KHz)   Window 8, 38 [LPC], Shift 2.5 (ms) Var.max 74.6, V/UV 0.16

J.3 Spectra in generated codebook

american

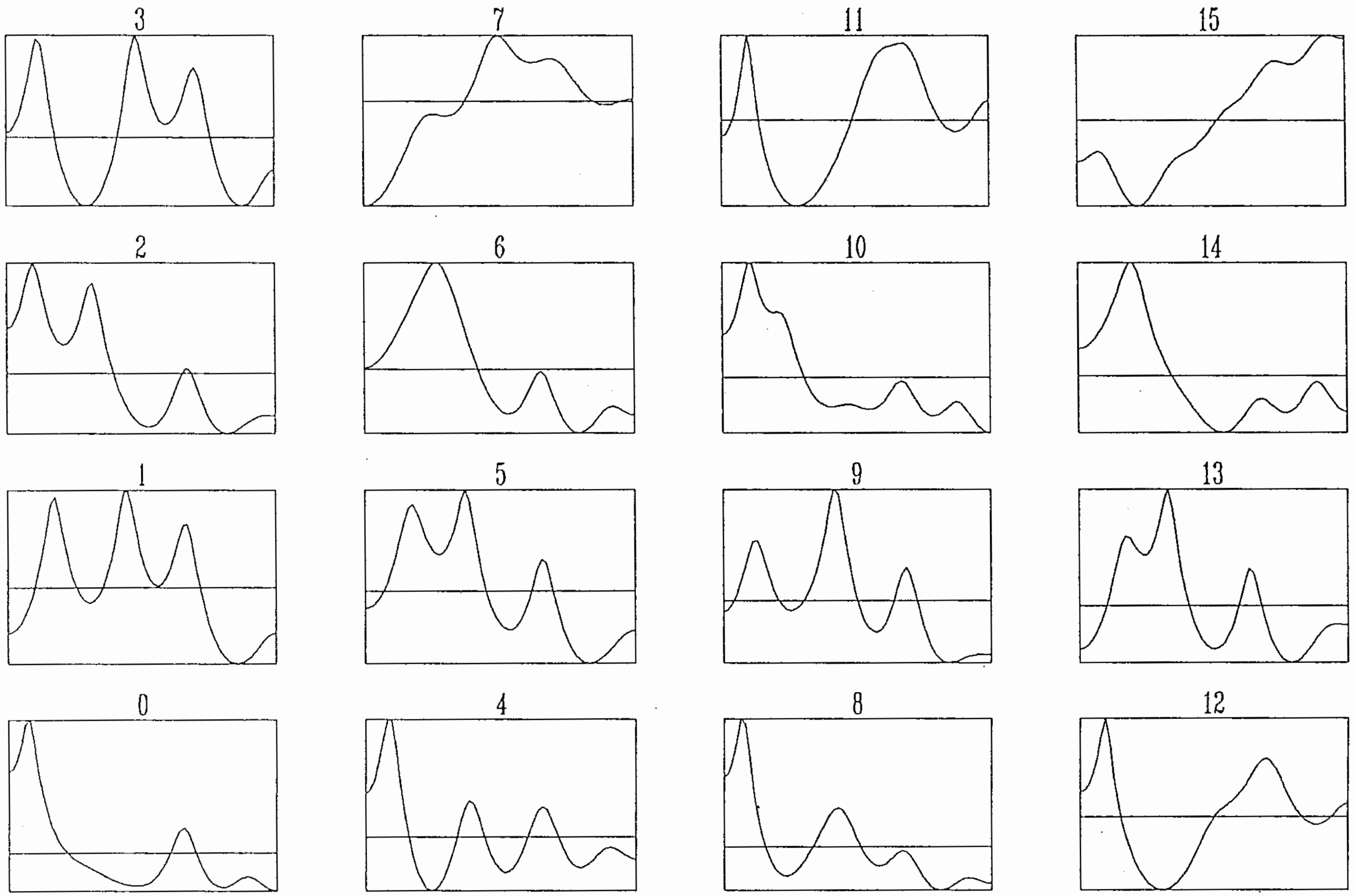
parameter	value
number of codes	16
distance	WLR
amount of test data	4
frame shift for training	128
frame shift for test	64



Spectra in Generated Codebook

parameter	value
number of codes	16
distance	WLR
amount of test data	4
frame shift for training	128
frame shift for test	64

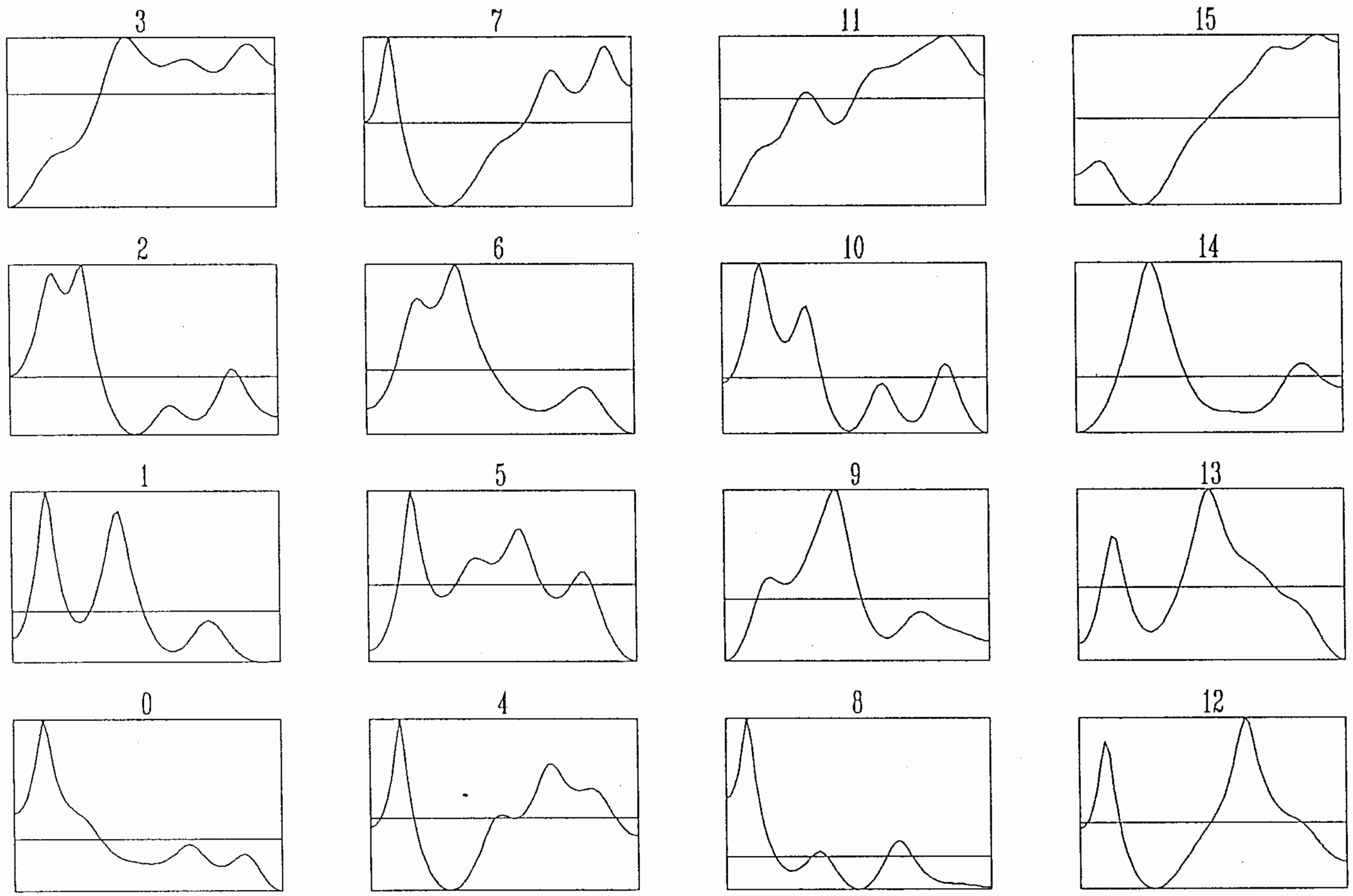
arabic



Spectra in Generated Codebook

chinese

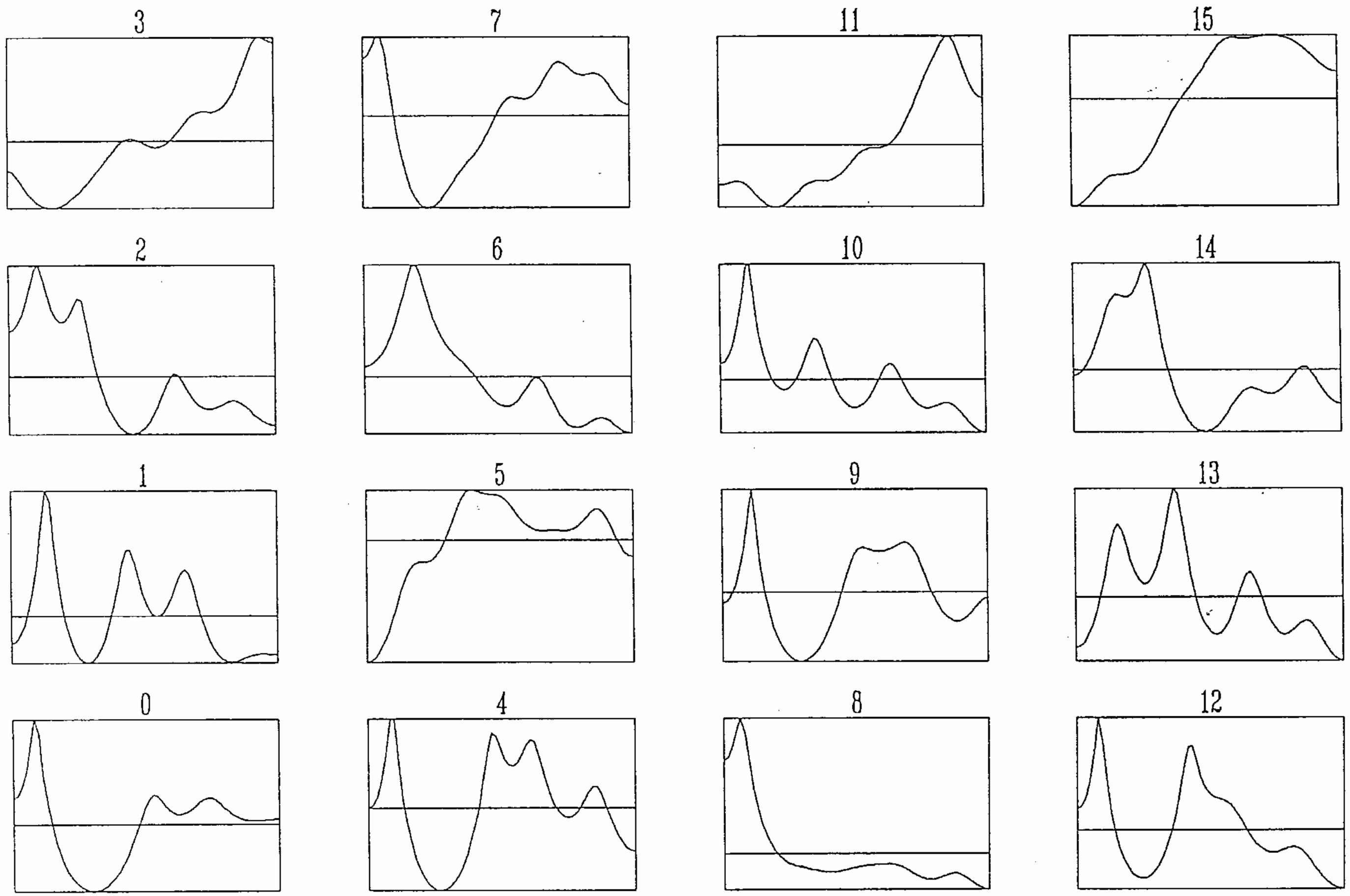
parameter	value
number of codes	16
distance	WLR
amount of test data	4
frame shift for training	128
frame shift for test	64



Spectra in Generated Codebook

french

parameter	value
number of codes	16
distance	WLR
amount of test data	4
frame shift for training	128
frame shift for test	64



Spectra in Generated Codebook