TR-I-0296

Speaker-Independent Features Extracted by a Neural Network and Their Evaluation in Speech Recognition

ニューラルネットワークを用いた話者に依存しない特徴の抽出と音声認識における評価

Ioana Donescu　　加藤 喜永　　杉山 雅英

1993年2月1日

内容梗概

本稿ではニューラルネットワークを用いた不特定話者の特徴抽出アルゴリズムについて述べる。本アルゴリズムにより任意の次数を持つ話者正規化特徴を取出すことができる。アルゴリズムは、(1) ファジィ級関数によって計算される教師を用いたニューラルネットワークの初期学習、(2) DTW による標準話者と新しい話者との整合、(3) ネットワークの追加学習の3つの処理からなる。ニューラルネットワークには、多入出力素子を持つ FPM を採用する。FPM の教師信号にはカテゴリ依存タイプとカテゴリ独立タイプの2種類を用いる。本アルゴリズムを FPM-HMM 音素認識と FPM-LR 文節認識により評価し、認識性能を従来の HMM や FPM と比較する。実験結果からニューラルネットワークが不特定話者の新しい特徴抽出器として使用できることを示す。

SPEAKER INDEPENDENT FEATURES EXTRACTED BY A NEURAL NETWORK AND THEIR EVALUATION IN SPEECH RECOGNITION

I. Donescu	Y. Kato	M. Sugiyama
INT	ATR	ATR

February 2, 1993

Abstract

In this report, an algorithm is proposed for the use of a neural network as a speaker independent feature extractor. This algorithm can extract normalized features with an arbitrary number of dimensions. In order to evaluate the performances of the proposed algorithm, a combination with continuous type HMMs, for several numbers of continuous density mixtures is tested. For comparison, several phrase recognition experimental results are given. The recognition rate is around 70%, but many directions are to be investigated in the close future. It is believed that a neural network can be used as a new speaker independent feature extractor and give good results, especially in language identification.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Foreword

The research project here presented covered a 6 months period, from August, 1 1992 to January, 29 1993. As the author was not familiar with the field of speech recognition, a month of theoretical study was required. Several very useful articles and book chapters, that will not be described here formed a very useful basis of study ([1], [6], [3], [2]).

As described in the abstract, the subject proposed combines neural network techniques -for feature extraction- and HMMs continuous density type models -for evaluation of the recognition performance.

1.2 Motivation of this study

Why the subject proposed was "Neural Network based Speaker Independent Features and Their Evaluation in SpeecH Recognition? Several points have to be pointed out, to explain the motivation of this study.

The key concept is *speaker-independent feature extraction*. Speaker independent speech recognition is a most up-to-date field investigated by researchers. Some important breakthroughs have been achieved recently. Most of the already tested systems use LPC cepstrum parameters, or FFT spectrum parameters. The fact is these parameters include several kinds of information, as speaker characteristics, emotional characteristics and other types of information. But they are *a priori* not decomposable into *phonemic* information and others. So, other ways to obtain speaker independent features - which would contain only phonemic information - is worth investigating. Or, a neural network - and more precisely the FPM (Fuzzy Partition Model) gave satisfactory results in speaker independent speech recognition([7]). Other examples of speaker-independent speech recognition using a neural network are : Feature mapping using a neural network[8], a method of constructing speaker-independent codebooks for discrete Hidden Markov Models (HMMs)[9], and other studies [10][11].

In [7], a 4-layer FPM is used as a phoneme classifier. The target is the recognition of 25 Japanese phonemes. The output layer contains only 1 unit, and 25 outputs. Each output corresponds to one phoneme and the network is trained for the activation of the corresponding output for each phoneme, independently of the speaker (pattern classification). The ideal target is when a phoneme is the input, the output contains 1 for the phoneme and 0 for the rest. From that results, it was assumed that the FPM was able to extract speaker-independent features. The paper to be presented by Kato in ICASSP, 1993, "Speaker-Independent Features Extracted by a Neural-Network" is the theoretical basis of this study. This paper proposes an algorithm for using a neural network to normalize features that differ between speakers.

After the theoretical formulation and description of a new feature extractor, another question raises. How to evaluate the method proposed?

That is the second point, the combination with the FPM with HMMs, as the neural networks are generally useful for pattern classification and HMMs are better for the treatment of the time variability of speech. Nevertheless, in our study we did not focus on the combination between FPM and HMMs, and the recognition results tend to prove our choice was probably not the best. This point will be discussed in chapter 5.

So, let us first describe theoretically the proposed algorithm, giving after several details about the achievements and experimental conditions. Finally, the experimental results, followed by a section comments and discussion.

Chapter 2

Formulation of a feature extraction NN

The fundamental idea is mapping, using a neural net, from aspeaker dependent speaker space (in our case the FFT spectrum analysis coefficients) to a speaker independent speaker space (space of the NN's outputs). The idea is so :

Speaker independent feature space
Arbitrary size of vectors
$\Uparrow$
$\boxed{\text{MAPPING}}$
$\Uparrow$
Speaker-dependent features

The second point is that the number of outputs is a free parameter, independent of the number of phoneme categories, unlike the conventional pattern classification.

Two questions are then raised.

- First, what target signal should be given ?

  This question will be largely discussed in the description of the algorithm.

- Second, what criterion should we use to designate the neural network?

Two solutions are proposed to each of these questions.

In the proposed algorithm, we distinguish 2 sets of speech data for 2 different uses : a *standard speaker* S is used for the target signal generation and the initial training. A set of several *arbitrary speakers* $\bigcup_A A$ is used for speaker independent training.

The criterion proposed is the minimization of the distance (for the metric chosen) between the sequence of target signals (calculated from the speech data of *the standard speaker*) and the sequence of actual outputs of the NN for any arbitrary input speaker.

2.1 The formulation

An algorithm is proposed to use a Neural Network as a feature extractor. The target is calculated using the acoustic patterns of one speaker, so called the *standard speaker*. The *standard speaker* is chose arbitrarily as no *a priori* rule or definition is given for "what is a *standard speaker*". As there is no human voice that can guarantee 100% of recognition, the random choice seems justified. Moreover, at the end of the procedure the neural network is supposed to be speaker independently trained, so the influence of the choice of the standard speaker should diminish.

The training of the neural network is performed twice : a first phases of initialization, speaker dependently, using the acoustic patterns of the *standard speaker* (U^S) and a second "incremental training" using a set of arbitrary speakers ($\bigcup_A U^A$).

2.1.1 Mathematically

Let f be the function of the NN, so f is only determined by the set of weights and F the target function. F is calculated for each speech segment for the standard speaker. Several choices are possible, but in our case we chose F to be the Fuzzy membership function.

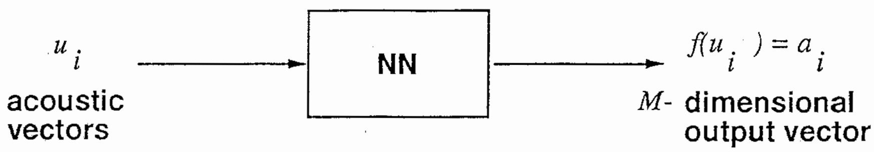


Figure 2.1: Concept of NN based feature extraction

The mathematical formulation is :

$$f : \bigcup_A U^A \longrightarrow P, \tag{2.1}$$

$$F : U^S \longrightarrow P, \tag{2.2}$$

We will note $F(u^S)$ the target vector calculated from the input *standard* vector $u^S \in U^S$ and $f(u^A)$ the actual output for a given vector in the arbitrary speaker's input space U^A.

P is the output space, and , according to the constraints of the FPM P is a hyperplane in R^M. That means that P is defined by the 2 constraints:

- Positive vectors :

  $\forall\ p \in P,\quad p = (p_m),\ \ m \in [1 \ldots M],\ \ p_m \geq 0$

- The sum of outputs is restricted to 1 :

$$\forall\ p \in P,\quad p = (p_m),\ \ m \in [1 \ldots M], \sum_1^M p_m = 1$$

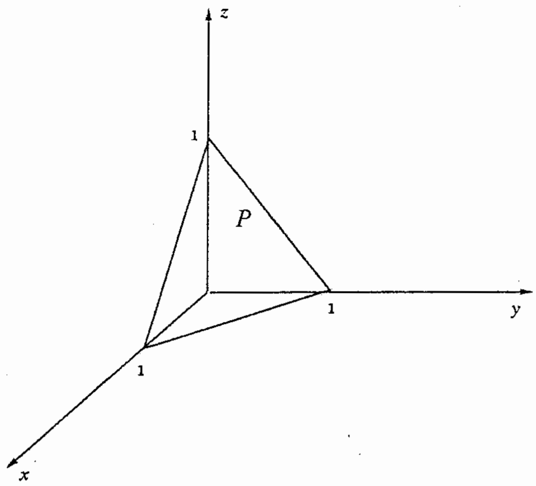


Figure 2.2: A 3-dimensional hyperplane: P

P is *the* speaker independent feature space to which the FPM performs the mapping. The constraints described above are also called "probabilistic constraints", characteristic of the FPM. These constraints allow the use of the Kullback divergence as a distance measure for the neural network.

These mathematical notations will be very useful in the following.

Let's give, before detailing, the 4 steps - that is to say the precise algorithm.

2.1.2 The precise algorithm

The 4 steps of the algorithm are :

1. Generation of the target F
2. Initial training using only the data for the standard speaker.
3. DP matching between the arbitrary speakers and the standard speaker - in order to decide the target signal for any input speaker.
4. Incremental training for several arbitrary speakers.

At the end of step 2, the neural network is speaker dependently trained, for the *standard speaker.* After step 3, a target signal is matched to any input speech sequence for any input arbitrary speaker. At the end the last stage, the neural network should be speaker independently trained and ready for the use as a **speaker independent** feature extractor.

2.2 Target generation

The target is calculated from the "phoneme" feature vectors of the standard speaker. That means the input data is a certain number of feature vectors for each of the K (K =25 in our case) Japanese phonemes. We'll speak about **phoneme categories**. We can consider that the set of segments (phoneme vectors here) for the standard speaker is divided in K subsets :

$$U^S = \{U_1^S, ..., U_K^S\} \ .$$

Figure 2.3: 25 Japanese phonemes corresponding to the network outputs

Phonemes	/b/,/g/,/d/,/p/,/t/,/k/,/m/,/n/,/N/,/s/,/sh/,/h/,/z/,/ch/ /ts/,/zh/,/r/,/w/,/a/,/i/, /u/, /e/, /o/, /silent/

As the target vectors are calculated using the fuzzy membership function (the calculation will be detailed in 2.2.12.2.2), a set of reference vectors - or templates - is needed. Which means that first we perform a VQ in order to obtain M representative vectors or templates, using a k-means clustering. This is a very important step, as there are 2 cases to be taken into consideration. There are 2 possible ways of generating these templates, dependent or independent of the phoneme category. In both cases, each target vector is calculated so as the sum of it's coefficients to be 1, which corresponds to the constraint on the FPM's outputs (cf FPM definition in appendix B).

2.2.1 Category-dependent target generation

In this cases, the templates are dependent on phoneme categories, and somehow characteristic of these phoneme categories. We consider, as mentioned above, 25 categories.

A set of M templates is generated from each category. So, at the issue of this VQ we have $25 \times M$ templates, which we write as :
$T = \{T_{ij}\},\ \ i \in [1, ..., 25],\ j \in [1, ..., M]$ the jth template of the ith category.

Using this set of templates, the target is calculated for each $(u^k)^S \in U^S$ by :

$$(a^k)^S = F((u^k)^S, T) \tag{2.3}$$

The result, $(a^k)^S$ is a 25-dimensional vector. In order to avoid too heavy notations, the S will be omitted.
$a^k = (a^k_n),\ \ for\ \ n \in [1, ..., 25]$ is given by :

$$a^k_n = \sum_{q=1}^{M} a^k_{nq} \tag{2.4}$$

$$a^k_{ij} = 1 \Big/ \sum_{l=1}^{25} \sum_{m=1}^{M} \left\{ \frac{||u^k - T_{ij}||}{||u^k - T_{lm}||} \right\}^{1/(g-1)}, \tag{2.5}$$

The distance $||\ \ ||$ is the Euclidean distance, and g represents the fuzziness. a^k_{ij} indicates the similarity between the template T_{ij} and the given vector u^k. If u^k is very close of T_{ij}, a^k_{ij} is close to 1, and vice-versa. The fuzziness coefficient g accentuates more or less the tendency to a "0/1" distribution. For a big fuzziness the output distribution tends to be quasi discrete. For example, if the fuzziness g is of 1.1, it means $1/(g-1) = 10$. If a given vector u^k is close to the template T_{ij}, $||u^k - T_{ij}||$ will be close to 0, and $\left\{ \frac{||u^k - T_{ij}||}{||u^k - T_{lm}||} \right\}^{1/(g-1)}$ will tend to 0. The sum will trend to 1, and the other coefficients will be very small as compared to this a^k_{ij}.

The summation is necessary inside each category, as the number of outputs desired is 25. The summation does not change the signification of each output coefficient a^k_n, which will represent the "similarity" between input vector u^k and the kth phoneme category. So, in the category dependent case, the FPM is not only a feature extractor, but also a phoneme classifier.

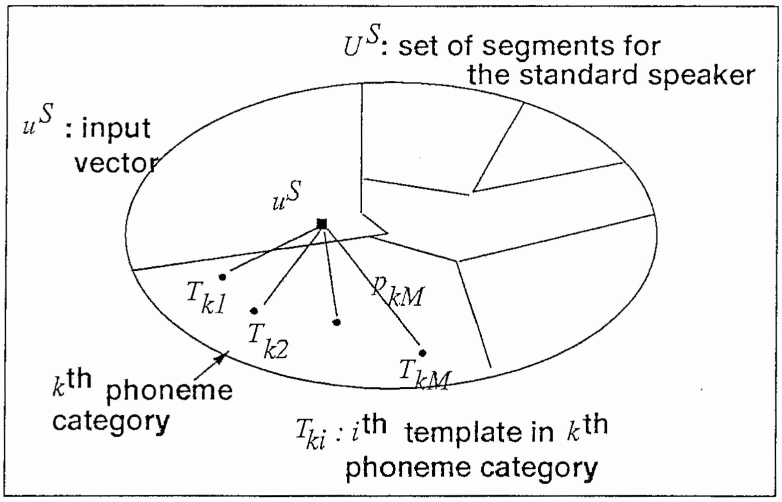


Figure 2.4: Category dependent target generation

2.2.2 Category-independent target generation

In this case, the 25 subsets are gathered together, so as the set U^S is one whole. Using a modified *k-means* clustering M templates T_m are calculated. These templates have no "physical" meaning, as the initial phonemic vectors are mixed. The number of outputs of the FPM is M in this case, and they are obtained more easily, by the following :

$$a_n^k = 1 \Big/ \sum_{p=1}^{M} \left\{ \frac{||u^k - T_n||}{||u^k - T_p||} \right\}^{1/(g-1)} , \tag{2.6}$$

The coefficient a_n^k still gives an evaluation of the similarity between u^k and T_n, but has no meaning in a phonetic sense. In this case, as there is no pattern classification, a combination with a phoneme classifier - HMMs in our case - is necessary.

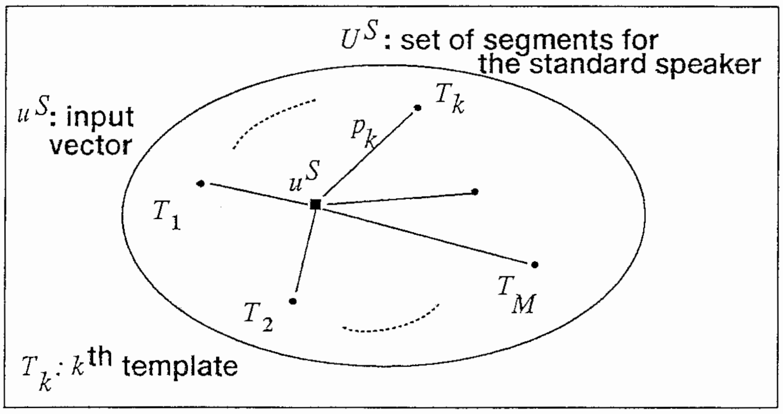


Figure 2.5: Category independent target generation

2.3 Initial Training

At this stage we have the pair $(U^S, F(U^S))$ for the initial training data (phoneme samples). The NN is trained by the criterion of the minimization for the total distortion. The measure of the distortion is the *Kullback* divergence (D - cf appendix C), as mentioned above (one of the particularities of FPM). A first approximation of the function of the NN, f is given :

$$\min_f \sum_k D(F((u^k)^S), f((u^k)^S)) \tag{2.7}$$

so the global minimization of the distortion between the input and the target. In fact, the learning is repeated until the global distortion is inferior to a threshold value ϵ. The first 2 steps are reminded by the figure 2.6.

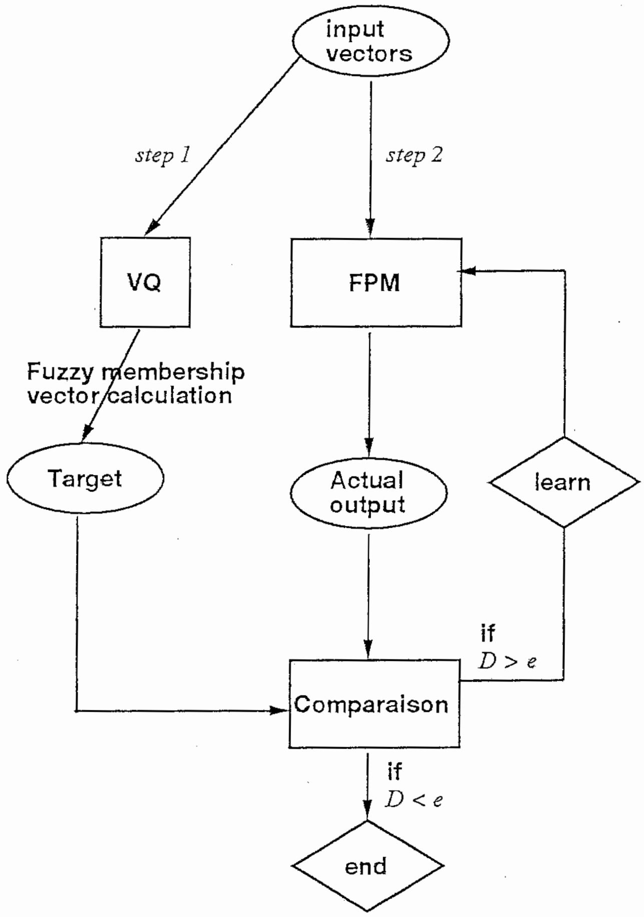


Figure 2.6: First 2 steps of the algorithm

2.4 Alignment between speakers

Till now the NN has been trained speaker dependently. Or we need speaker independent features. So, we use several sets of speech segments ($\bigcup_A U^A$) for a certain number of *arbitrary speakers.* In this stage, the segments considered are words. The standard speaker is included in this set. The data used is no longer already labeled, so the data will be automatically segmented from words to phonemes. The aim of this step is to give at target for each data vector input. The precise procedure is described in Fig. 2.7.

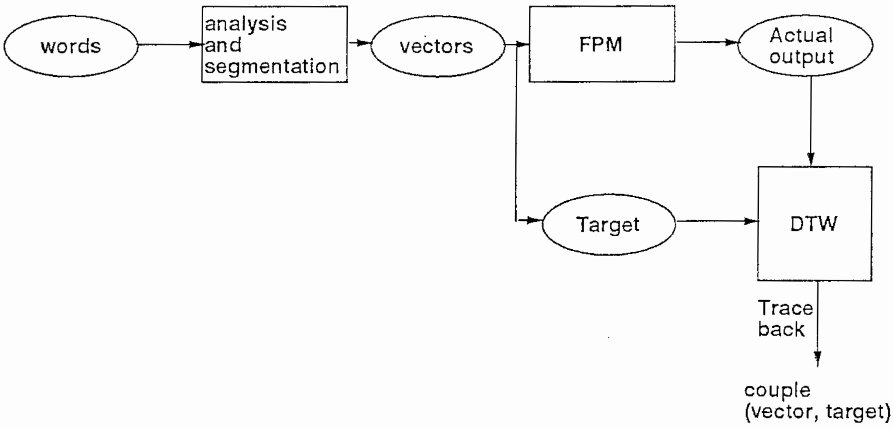


Figure 2.7: Automatic alignment procedure

In the DTW, the distance was calculated by the *Kullback* divergence as well. The idea is to find the minimal cost path of alignment between the target segment -calculated for the standard speaker, using the templates calculated in the previous step and the function described in 2.2.2- and the FPM output segment.

Let us suppose, for a same word W:
$W^S = (u_1^S, ..., u_I^S)$
$W^A = (u_1^A .., u_J^a)$
the sequences of input vectors, for the standard speaker S and the arbitrary speaker A, I and J being the respective sequence lengths.

The alignment is given by the following formula :

$$D(\boldsymbol{W}_q^A, \boldsymbol{W}_q^G) = \min_{\sigma,\tau} \sum_{i=1}^{I+J} d(F(u^A_{\sigma(i)}), f(u^G_{\tau(i)})). \tag{2.8}$$

At the end of this step, by tracing back a target vector is given for each input vector, for each word, each speaker. The next and last step is the incremental training.

2.5 Incremental training

The procedure is the same as for the initial training, only the data is changed (see FIG. 2.6). So, the function f is definitely adjusted and the FPM is trained.

The criterion is summarized by :

$$\boxed{\min_f \min_{\sigma,\tau} \sum_{u^A_{\tau(i)} \in U^A} D(F(u^S_{\sigma(i)}), f(u^A_{\tau(i)}))} \tag{2.9}$$

At this stage, we have to precise that a neural network trained by the algorithm described above will be called "alignment-based", by opposition with the classical pattern classifier - so-called "label-based".

For reference, in the case of the "label-based" classifier, the criterion is :

$$\min_f \sum_{u_i^A \in U^A} D(F(u_i^A), f(u_i^A)) \tag{2.10}$$

as there is no need of a standard speaker S and no use of time-alignment functions. $F(u_i^A)$ is simply given by, if $u_i^A i$ is in phoneme category k :

$$\begin{array}{ll} & 1 \ldots \mathrm{k} \ldots \mathrm{M} : \text{phoneme label} \\ F(u_i^A) = & (\, 0 \ldots 1 \ldots 0) \end{array}$$

In order to test the robustness of the features extracted, in the experiments we will use a completely different set of data (speakers and words uttered). The results in continuous speech recognition, using this method combined with an LR parser gave encouraging results ([14]). In the present study, in order to evaluate these features, we combined the FPM with HMMs.

After this more theoretical and general view, let's give details about the present study.

Chapter 3

Experimental procedure

The first idea was the evaluation, using a combination with HMMs of the FPM as a feature extractor, in the category independent case, and a comparison with the performance of the "category-dependent" FPM, which was already programmed. The first work was to adapt the category-dependent to the category-independent, for an arbitrary number of templates, which required dynamic programming. Unfortunately, memory allocations sometimes need more memory than available and slow down the execution.

The global procedure is given in Fig. 3.

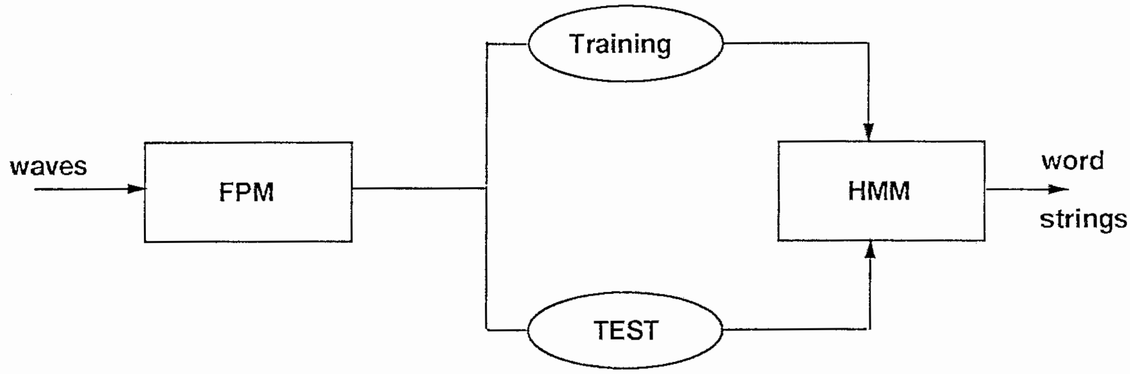


Figure 3.1: Phoneme recognition system

The FPM box implies a NN already trained. In fact, the most important part of this work was the NN training. For the HMM part, HTK - **HMMToolKit** ([5]). The use of this already programmed package largely facilitated this work (see chapter 3.3).

3.1 Category independent neural network training

3.1.1 The architecture

The architecture for the neural network is the same for both category-dependent (CD) and category-independent cases (CI). A four layers, feed-forward neural network.

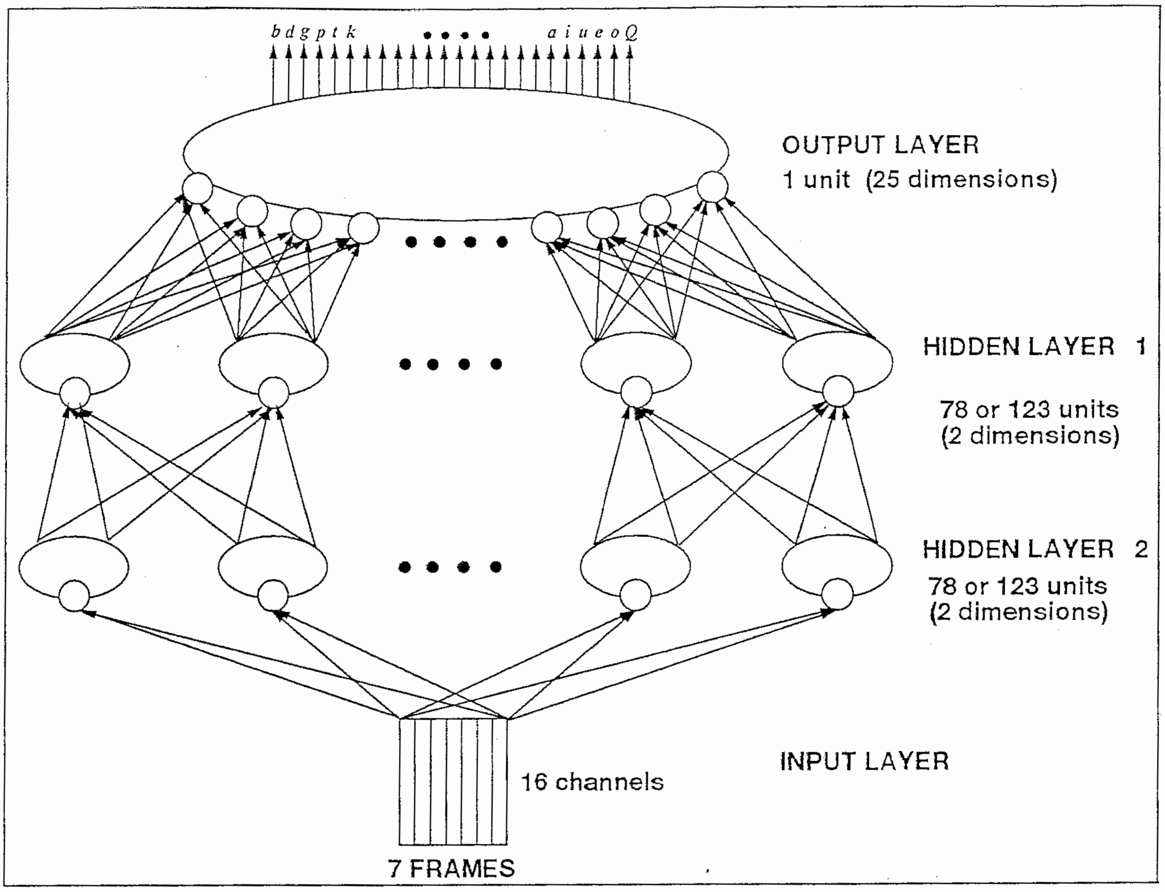


Figure 3.2: FPM architecture for Japanese 25 phoneme recognition

The input and the output layer have have only 1 unit. For the input layer, this unit has 113 dimensions, because the input vectors are 112 dimensional (due to the analysis, 16-channel mel-scaled spectrum with 7 frames). The size of the output is 25. We chosen 25 as well for M in the CI case, taking into consideration comparison reasons and computational cost. The 2 hidden layers are composed of 123 units, 2 dimensions. This values was chosen in order to compare FPM performances to TDNN. The number of free weighted parameters in TDNN is 24825, slightly higher than the 24648 in the FPM with 78 units in each hidden layer. Using 123 units in the hidden layers, we need 50307 free connections, that is to say twice more than for TDNN.

3.1.2 Speech database

The data used for the FPMs training was taken from a Large Vocabulary database of 5240 Japanese words uttered by 8 native Japanese male speakers. The words were picked up so as to have about 2000 samples for each phoneme. The data used for training and testing the HMMs was taken for a 216 balanced Japanese words, 10 male speakers database. 8 speakers have been used for training and 2 for the tests.

For the initial training, we used 2620 words from the Large vocabulary Database, about 50000 vectors, that is to say around 2000 vectors per phoneme.

3.2 Experimental steps

Let's describe the sequence of experimental steps, corresponding to the theory explained above.

3.2.1 VQ and target calculation

The first problem was to reduce the quantity of the data : a *k-means* on a space containing 50000 vectors of size 112 was impossible to realize. So, we added a preliminary step : using the LBG algorithm, each phoneme subset was reduced to 128 vectors. The next step is a modified k-means algorithm to output the 25 templates. Here is the detail of the algorithm, where the current number of templates is L, and the index of the cluster to be splitted K.

1. Initialization : Initial template (T_0) is chosen randomly.
   L is set to 1 and K to 0.
2. Choice of a new initial template T_L: Search of the vector u_i so as :

   $||T_L - u_i^K|| = \max_j ||T_0 - u_j^K||$
   $T_L = u_i$
3. Splitting U_K into 2 clusters using k-means algorithm, using the initial templates T_L and T_K. Set L to $L+1$
4. If $L=$ desired number of templates (25 in this case) finish.
   Else go to step 5.
5. Search the cluster of maximal distortion. Let's d_i be the total distortion for cluster i. Chose p as : $d_p = \max_{i=1}^{L}(d_i)$. Set $K = p$. Goto step 2.

Using this algorithm, the total distortion decreases as expected. See Fig. 3.2.1. In this case, the distortion is measured by the global euclidian distortion on the "whole" space U^S.

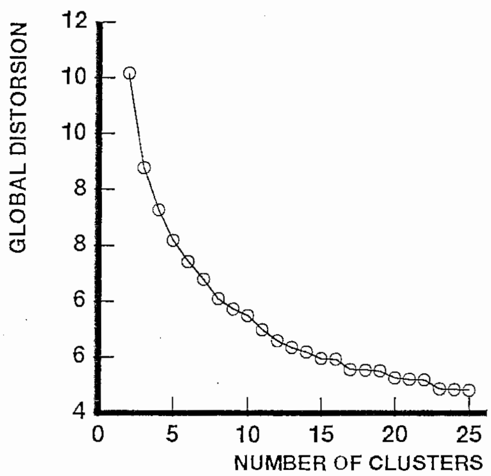


Figure 3.3: Evolution of the total distortion during the VQ

The decrease continues when increasing the number of templates. Nevertheless, it was difficult to carry out the experiments for more than 25 templates. Indeed, we could not handle more than 25 outputs for the FPM.

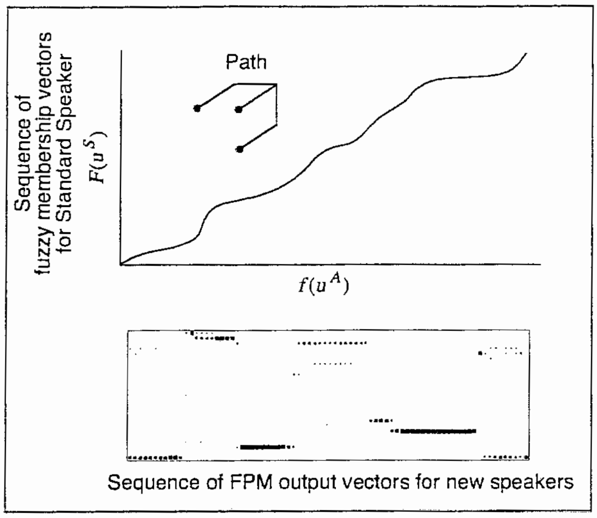


Figure 3.4: Time alignment procedure

The next step was the calculation of the target vectors, already described. The fuzziness value g is 1.1, which implies : $1/(g-1) = 10$. The fuzziness value is discriminative.

3.2.2 Initial training

Before presenting the "phoneme" samples to the FPM, with their corresponding target, we mixed this data randomly, so as not to focus on a particular phoneme category, and then completely change all the weight parameters for another one (imagine presenting first all vectors for /b/, after all vectors for /d/ and so on). The distortion was expected to decrease in the same manner. Nevertheless, several tries were necessary to find a suitable value for the rate η (B). The final choice for η was 0.005. Here is the evolution of the total distortion. The initial training was stopped after 63 iterations. The total distortion - in this case average Kullback divergence per sample - was of 0.011375.

3.2.3 Time-alignment and pattern matching

This stage is the data preparation for the incremental training(see Fig. 2.7). We use 8 speakers including the *standard speaker*.

Table 3.1: Training speakers

MHT, MMS, MMY, MSH,
MTK, MTM, MTT, MXM

The data segments (words) are presented to the FPM in this order. The first speaker presented is the standard speaker, for which, after analysis, the target signal is calculated. For the 7 following segments, the FPM outputs are calculated only. Then, the DTW is performed as shown in Fig. 3.4, between the sequence of target vectors for the given segment and the sequence of actual outputs of the FPM.

The path is a symetric path, given by :

$$G(i,j) = \min \begin{bmatrix} D(i,j) + 2D(i-1,j) + G(i-2,j) \\ 2D(i,j) + G(i-1,j-1) \\ D(i,j) + 2D(i,j-1) + G(i-1,j-1) \end{bmatrix} \tag{3.1}$$

where G is the minimal cost function and D is the Kullback divergence (see appendix C). So, for any input speaker, the target is aligned on the target for the same word for the standard speaker, dynamically. The couples (vector, target) are obtained by tracing back. The can be mixed (in order to have a random distribution) or kept in the order of the speech segments, as there is no phonemic classification.

3.2.4 Incremental training

Given the set $(U^A, F(U^A))$ obtained in the previous step, there's no difference between the incremental training as compared to the initial training. The "initial" weight file is obviously the last weight file (last iteration) in initial training. The global divergence -calculated as an average per speech sample- is much higher, and the learning slower. The evolution of the total divergence is given in Fig.

Once the *category-independent* neural network was trained, the next step is the evaluation, by combination with HMMs. A *category-dependent* (8 and 16 templates per phoneme category) already programmed, has been tested in the same conditions. The precise procedure for the training and test of the HMMs is described next.

3.3 Recognition performance using HTK software

HTK is a package developed at Cambridge University which deals with continuous density mixture HMMs. Any number of states, mixtures, any vector size and full or diagonal covariance matrices are allowed.

In our case, the size of the observation vectors was 25, and the output files of the FPM have been written in HTK format. The HTK format consists in a continuous sequence of sample vectors preceded by a header.

The format of this header is :

Num Samples:number of samples in file	Sample period: 10000.0 usecs
Sample Size: (in bytes)	Sample Kind: WAVEFORM_E

- The unit *usecs* is 1 usec= 100ns.
- WAVEFORM_E : $sampled \quad waveform + energy$
- The number of samples in file differs from word to word.
- The sample size, for a 25 dimensional vector is of 100 bytes.

In parallel, the corresponding "label files" are necessary. They contain transcription information, consisting of a sequence of labels (in our case words) with their start and end points.

For example, the word "poketto" (cf. the english origin "pocket" → MHT_B_0020.lab) is given by :

```
0 1250000 sil
1250000 1400000 p1
1400000 2400000 o
2400000 3200000 k2
3200000 4150000 e
4150000 7050000 tt
7050000 8150000 o
8150000 9310000 sil
```

3.4 Basic Tools employed

The 4 basic functions employed all the long of this study, for the creation, training and evaluation of each HMM model are the following.

First, the model is created using *HMakeModel* - not a HTK function, but a function written by D.Rainton for the proper use in ATR. This function makes HMM model in HTK format, left-to-right, with either 0 or single state skipping.

Next, the model is **initialized** - estimates of initial means and variances using a set of observation (=training data) sequences, using *HInit* function. In our case the observation sequences are words. *HInit* provides general estimates for the phoneme seed models by cutting from the training words the data corresponding to the labels. It works by using Viterbi alignment to segment the training vectors, and then recompute means and variances. In the multiple density mixtures case, vector pools are clustered using a K-means algorithm.

The output of *HInit* is normally input to *HRest*, which performs the Baum-Welch **reestimation** of the parameters of a single HMM model, using the same set of sequences (=training data). *HRest* as well cuts out the phoneme data from the training segments automatically. At the end of this stage the "HMM training" is accomplished.

The next step is the **recognition** step. Given the test data, another HTK basic function, *HVite* is used as a Viterbi decoder with finite state syntactic constraints. In our case the grammar is composed by the sequence of phoneme model, and no constraint is imposed (see table 3.2). So, the matching is done phoneme after phoneme, any phoneme being allowed to follow any other. The Viterbi recognizer will match test data files against the "network" of phonemes and output a transcription file in HTK label file format.

Table 3.2: The "grammar": phoneme network in HTK

```
$phn = a | i | u | e | o | p1 | p2 | t1 | t2 | k1 | k2 | b1 | b2 | d1 | d2 | g1 | ng | m | n |
N | r | w | y | s | sh | h | z | ch1 | ch2 | ts1 | ts2 | sy | hy | zy | cy | py | ky | by | gy1 |
ngy | my | ny | ry | aa | ii | uu | eei | oou | sil;
(<$phn>)
```

Finally, the last step is the output of the recognition results, using *HResults*. This function reads the HTK format files output of *HVite* and compares them with the corresponding transcription files by dynamic programming. The basic format of the output files of HResults is the following:

```
------------------------ Overall Results ------------------------
PHRASE: %Correct=0.54 [H=2, S=366, N=368]
PHONE:  %Corr=66.76, Acc=50.27 [H=1984, D=8, S=902, I=490, N=2972]
-----------------------------------------------------------------
```

The first line gives the number of label files identical to the transcription files (in our case, it would be a "word recognition" performance). The second line gives the **phoneme recognition rate**. The results are based on DP matching.

- H : number of correct labels
- D : number of deletions
- S : number of suppressions
- I : number of insertions
- N : total number of labels in the defining transcription files

The correction percentage is calculated by : $\%Corr = H/N * 100$ and the accuracy : $Acc = \frac{H-I}{N} * 100$

An extra output available is the confusion matrix. If the are any unknown label in the transcription files (labels for which there is not enough data to calculate a HMM model, N-ny or double consonants : kk, tt, pp) they can be ignored in the matching process. Thus the "unknown labels" exist in the transcription files, the perfect recognition rate would still not be 100%.

Chapter 4

Experimental results

4.1 Reminder : Experimental conditions

As reminded before, the NN used was a 4-layer FPM, and the experiments were based on the FPM-HMM combination. The fuzziness was of 1.1 for most of the experiments. The analysis conditions are the standard ATR analysis condition (see Table 4.1).

Table 4.1: Experimental specification

Speakers	
Native male Japanese	
Number of speakers	Initial training: 1 Incremental training: 8
Acoustic Analysis	
Input pattern	16-channel FFT mel-scaled spectrum with 7 frames (70ms)
Frequency	12kHz
Frame rate	10ms
Window	256-point (21.3ms) Hamming
Power normalization	Normalized between 0.0 and 1.0 with the average at 0.5

For the FPM, the architecture shown in Fig. 3.2 was used for all the experiments, regardless of whether the target generation was dependent or independent of phoneme categories. In the category independent case, the choice is justified by comparison requirements and computational cost. After combining with continuous type HMMs, the recognition rate is output for 2 open speakers, i.e. speaker used neither for FPM training, nor for HMM training. These 2 speakers are MAU and MNM. Table 4.2 shows the experimental conditions for the HMMs.

Several questions can be raised :

- Is the NN really speaker-independently trained ? This means the need to calculate the alignment distortion after training.
- What is the recognition performance, how does it evolute with the number of Gaussian mixtures ?
- Is this method better or worse than the classical feature extraction (LPC for example) combined with HMMs ?

- What is the recognition performance for the *standard speaker*, using the target signal as an input for the HMMs ? Could we find a "best" value for the fuzziness and in this way optimize the target generation ?
- Finally, is there any mean to improve the target calculation in the category dependent case without increasing the number of outputs ?

Table 4.2: Experimental conditions of HMM phoneme recognition

Task	49-phoneme recognition
Speakers	Native male Japanese Training: 8 Open: 2
Input parameter	25-dimensional FPM output vector
HMMs	Continuous mixture density type 4-state, 3-loop 49-phoneme models
Phonemes	48 phonemes + silent class */p1, p2, t1, t2, k1, k2, b1, b2, d1, d2, g, ng, m, n, N, r, w, y, s, sh, h, z, ch1, ch2, ts1, ts2, sy, hy, zy, cy, py, ky, by, gy, ngy, my, ny, ry, aa, ii, uu, eei, oou, a, i, u, e, o, silence/*
Training data	1,728 balanced Japanese words (216 × 8 speakers)
Open data	434 balanced Japanese words (216 × 2 speakers)

4.2 Results

Before giving the recognition performances, let's give some results in terms of alignment distortion.

4.2.1 Alignment distortion and target generation

The influence of target generation has been evaluated by calculating the average alignment distortion for 10 speakers, 8 training speakers and 2 open speakers, using 100 words per speaker. These words are exactly the 100 first words of the Large Vocabulary Japanese Database, that is to say most of them were already included in the phase of incremental training. Nevertheless, the 2 open speakers, that we noted M9 and M10 were not included in the training set. The average is calculated per speaker and per frame, and the distortion is given by the Kullback divergence. Table 4.3 shows that the incremental training makes the distortion uniform for all speakers in both cases. The distortion is slightly higher in the CI case for the open speakers, and the differences between speakers are emphasized. In the CI case, the results given in Table 4.3 are obtained using the NN tested by combination with HMMs.

Table 4.3: Alignment distortion for 10 speakers

Speaker	Distortion	
	Category-dep. (CD)	Category-indep. (CI)
M1	0.257	0.166
M2	0.338	0.341
M3	0.357	0.319
M4	0.357	0.470
M5	0.365	0.330
M6	0.351	0.347
M7	0.333	0.345
M8	0.342	0.326
M9	0.369	0.431
M10	0.398	0.484
Training	0.338	0.330
Open	0.384	0.457

Training Speakers: M1–M8
Open Speakers: M9, M10

One question is the relation between distortion for open speakers and recognition results? It seems that a higher distortion would lead to worse recognition results. But there is no assertion, because the set of data tested is a different one.

In order to compare the evolution of this distortion with the number of iterations, Table 4.4 are given. Unfortunately, by lack of time, there is no corresponding recognition result. It seems that the CI is still strongly dependent on the standard speaker, but this tendency diminishes with the number of iterations. After 100 iterations, the distortion seems to be completely equalized, and that the NN should be speaker independent. Unfortunately, because a lack of time the features extracted by this NN have not been tested. The discrepancies for the speakers who are far is increased (cf M4, M9, M10). Somehow, an average distortion around 0.3, almost constant for the majority.

Table 4.4: Alignment distortion for 10 speakers (CI case, several iterations)

Speaker	Distortion		
	10 iterations	50 iterations	100 iterations
M1	0.158	0.172	0.324
M2	0.357	0.352	0.326
M3	0.337	0.328	0.324
M4	0.454	0.518	0.324
M5	0.316	0.335	0.324
M6	0.342	0.363	0.326
M7	0.352	0.357	0.327
M8	0.313	0.344	0.324
M9	0.426	0.454	0.329
M10	0.464	0.506	0.324
Training	0.328	0.346	0.325
Open	0.445	0.480	0.326

Training Speakers: M1–M8
Open Speakers: M9, M10

4.3 Recognition performances

We tested, by combination with HMM, 4 types of FPM feature extractors:

- LB : conventional label-based NN
- AB/CD, 8 templates : alignment based, *category dependent*, 8 templates per category
- AB/CD, 16 templates
- AB/CI, 25 templates : *category independent*, 25 templates

Several numbers of Gaussian density mixtures have been tested.

Table 4.5: Phoneme recognition results with various number of mixtures

Number of mixtures	Network Type	Mistaken phonemes/2373 Insert	Delet	Subst	Recognition Rate	Accuracy
4	LB	612	304	674	58.79	33.00
	AB	376	274	600	63.17	47.32
5	LB	604	305	610	61.44	35.99
	AB	388	245	627	63.25	46.90
6	LB	561	291	573	63.59	39.95
	AB	378	266	545	65.82	49.89
8	LB	597	314	581	62.28	37.13
	AB	329	252	528	67.13	53.27
12	LB	616	312	557	63.38	37.42
	AB	337	227	491	69.74	55.54
16	LB	589	272	591	63.63	38.81
	AB	354	243	469	70.00	55.08

AB : 8 templates/class
LB : conventional training
Fuzziness : 1.1

The exact number of phonemes in the test data used in the experiments is given in Appendix, Table D.1. In fact, 2 slightly different sets of data have been used for the tests. In Table 4.5, all the 216 balanced words were tested (cf. Table D.1). But several transcription files contain "unknown" labels, like /N-ny/, /pp/, /tt/ and others. They have been ignored in the matching process. One solution to avoid the problem of transcription files containing "unknown labels" was to delete from the test data base all the files containing other labels than the modelised ones. The quantity of remaining data was:

- 188 words for MAU out of 216 balanced words.
- 180 words for MNM out of 216 balanced words.

The phonemic details are given in Table D.2 in Appendix.

Recognition performance for CD, 16 templates

Fuzziness : 1.1 Templates : 16 per phoneme category

Table 4.6: Phoneme recognition performance (CD, g=1.1, 16 templates)

Nb of mixtures	Recognition rate	Accuracy
4	76.11	62.45
6	76.82	62.18
8	78.77	65.21
12	79.71	67.43
16	79.41	67.53

These are the best results obtain for speaker independent experiments, in terms of "recognition rate". They show the influence of the initial clustering. The recognition rate is higher for a larger number of templates. These results also confirm the distortion for open speakers given in Table 4.3.

Recognition performance for CI, 25 templates

Fuzziness : 1.1 Templates : 25

Table 4.7: Phoneme recognition performance (CI, $g = 1.1$, 25 templates)

Nb of mixtures	Recognition rate	Accuracy
4	66.76	50.27
6	68.44	45.05
8	70.29	48.62
12	71.67	50.27
16	71.47	55.85

The corresponding confusion matrices are in Appendix E.

Comparison: FPM-based / LPC-based feature performances

In order to obtain a set of "reference" results, phoneme recognition rates for LPC-based features, under the same experimental conditions for the recognition part, have been evaluated. The analysis conditions for the LPC-based feature extraction are shown in Table 4.8. A comparison between the best results obtained by the method proposed in this paper and LPC-based results is shown in Table 4.9. The phoneme recognition rate is higher for LPC, but the data compression when using FPM extracted features has to be taken into account: 25 dimensional vectors instead of 34 dimensional vectors for LPC.

The number of templates is another important factor for recognition performance. It was shown that at least 25 templates per class are required for speaker-independent phoneme recognition. The AB/CD-FPM performance therefore can be improved using a larger number of templates. Memory cost for the neural network-based feature extractor is fixed, even with increasing number of templates, because the cost depends on the size (the number of weight parameters) of the network. Assuming 25 templates per class and the network size shown in Fig. 3.2, the network requires about 7/10 the cost of the LVQ algorithm[17].

Table 4.8: Analysis conditions for LPC-based feature parameter

Analysis order	14th order
Frequency	12kHz
Frame rate	5ms
Window	256-point (21.3ms) Hamming
Feature parameter	16th LPC cepstrum + 16th Δcepstrum + Power + Δpower

Table 4.9: Recognition comparison with LPC-HMMs and FPM-HMMs

MIXTURES	SAMPLE KIND	Phoneme recognition rate(%)
6	LPC	83.34
	Label-based	64.50
	Alignment-based	76.82
8	LPC	83.71
	Label-based	63.08
	Alignment-based	78.77
16	LPC	80.96
	Label-based	63.11
	Alignment-based	79.41

For reference as well, even if these experiments have not been carried out in this study, the comparison, in phrase recognition experiments, between the performances of label based and alignment based, category dependent FPM. These experiments have been obtained by combination with a LR parser with 1672 rules. The task was the recognition of 278 phrases.

Table 4.10: Phrase recognition results

Alignment-based	68.00
Label-based	72.00

4.4 Some results on the *standard speaker*

Another question which can be raised is "what is the recognition performance for the *standard speaker* ?"

The data used was the large vocabulary database (5240 words) and the 216 balanced Japanese words. The speech samples were separated in 2 categories :

1. 2692 words for training ("odd" words)
2. 2693 words for recognition ("even" words)

As the quantity of the database are not the same, the results are not directly comparable with the previous recognition rate results. To give a better basis of comparison, the recognition rate per phoneme and the confusion matrices are given in Appendix E, D.

4.4.1 Global recognition rate for $Fuzziness = 1.1$

The template file used is the same that for the CI, 25 templates experiments.

Table 4.11: Recognition rate for the standard speaker, $g = 1.1$, M=25

Mixtures	Recognition rate	Accuracy
4	74.49	58.97
8	79.33	63.38
12	81.18	68.02

4.4.2 Global recognition rate for Fuzziness=1.6

Results for the standard speaker, in the same experimental conditions, same template file, but for a higher value for the fuzziness.

Table 4.12: Recognition rate for the standard speaker (g=1.6, M=25)

Mixtures	Recognition rate	Accuracy
4	81.53	67.05
8	84.35	68.60
12	85.79	70.76

As the experimental conditions were the same, these results show that a better recognition performance is obtained when the fuzziness is increased, so the output vectors of the FPM are less close to a "0-1" distribution, but they are more smoothly distributed. In Appendix D, we give more precise results, including the recognition rate per phoneme sample, and the number of samples considered.

Chapter 5

Comments and discussion

The previous results have partly already been commented. The recognition performance is not very high as compared to the classical LPC cepstrum/HMM performance. But the data is compressed. And, as the recognition results for the *standard speaker* with a fuzziness of 1.6 shows, the optimal value for the fuzziness was not employed from the beginning. Given the recognition results for 1.6, and the alignment distortion for the open speakers after 100 iterations of incremental training, the recognition rate might improve, and for a lower number of mixtures. Unfortunately, 2 results are not yet enough to decide for the optimal value of the fuzziness. By the lack of time, other values have not been tested, so it is difficult to give a definitive conclusion.

The detailed results for the standard speaker show how difficult it is to generalize a recognition performance for 49 phoneme model, given the difference of proportion of sample, as well in the training as in the test data, for each phoneme. In some cases (see /p1/, /ch1/) the results are completely irrelevant, as 1/3 gives a 33.33% recognition rate. The results can only be compared in the same conditions, but not general conclusion can be drawn before a statistical study.

Given the results for the standard speaker, and the supposition that if the recognition rate for the standard speaker is drastically improved, the target calculation is more accurate, one interesting question is what happens if we try to reduce the size of the target ? The computational advantages are obvious, and as we mentioned, this was one of the advantages of this method as compared with the LPC cepstrum HMM. But, as the distortion calculated in the first step quantitizes the quality of the templates, the question is "How to reduce the size of the target vectors without a too big loss in the representativeness of the templates?" One possible answer is to include another step in the template generation (**Tree-based target generation technique**), in the category independent case. Let's give some details : after obtaining the set $T = (T_1, \ldots, T_M)$, let's perform a *k-means* in each of the subspaces of those templates. In fact, a "template-category" dependent clustering, and output L templates for each of the M subspaces shown in Fig.5.1. This calculation should be close to a clustering in $L * M$ templates, but the size of the output remains M. In this case, M can be diminished. The fuzzy membership function is calculated using these templates, in the same manner as for the category-dependent case.

One experiment have been carried out :

M= 10
L= 5
Fuzziness=1.6

The data for the standard speaker, in the initializing stage, is the same. The global distortion after clustering (the template file contains 50 templates) is of : 4.104761 as in the

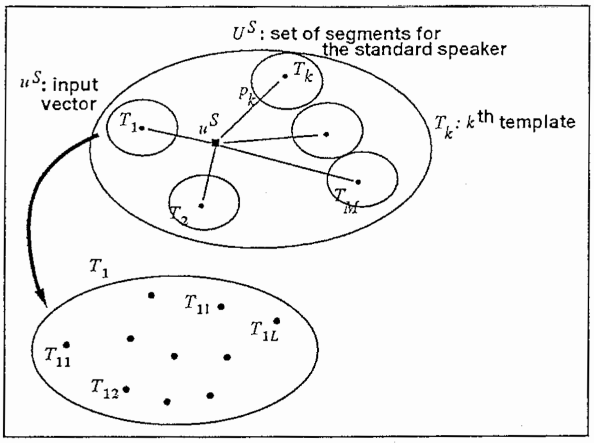


Figure 5.1: Tree-based target generation

case of $M = 50$ it is of 4.144501.

The recognition performance have been calculated for 4 mixtures, and the results is shown in Table 5.1.

Table 5.1: Recognition rate for the standard speaker using tree-based target generation technique ($g = 1.6, M = 10, L = 5$)

Mixtures	Recognition rate	Accuracy
4	82.33	68.77
4*	81.53	67.05

*: $g = 1.6, M = 25$

Another very interesting and unfortunately not investigated point is the recognition performance for the *standard speaker*, for a fuzziness of 1.6 and only 10 templates. The input vectors for the HMMs are only 10 dimensional, and the number of mixtures is only 4. Given the size of the vectors, it is probably better if the model has a low number of Gaussian mixtures. The possibilities are though very large, as they are many combinations possible for 2 steps clustering. If the results obtained for the standard speaker are confirmed by the recognition performance for the open speakers, the data compression achieved is very interesting (as the computational cost for the neural network is fixed once the architecture is decided).

5.1 Conclusion and future work

An algorithm for using a neural network was theoretically described, and some experimental results have been shown. These results show the effectiveness of the use of a neural network as a speaker independent feature extractor, even if the recognition performance obtained in the first experiments was slightly low - 70%. But the following experiments showed that at first neither the NN, nor the combination NN/HMMs have been optimized. We can consider a modular optimization - for example optimize the target signal generation, by modifying the value of the fuzziness and the size of the target vector. The results for the standard speaker show a good recognition performance for small size speech samples and low number of Gaussian density mixtures. Other types of improvements can be expected, as the fuzzy membership function for the target calculation might not be the best choice.

5.2 Future research

The field is widely open for future research in this study. Except from the optimization of the target signal generation, another possible direction to be investigated is the use of a global optimization criterion, like the minimum error classifier (MCE). A global approach is expected to provide much better results for a given task. Another possible application, except from phoneme recognition or continuous speech recognition, is language identification. Language identification is a very interesting study and was proposed as a topic for this 6 months project. Unfortunately, the author couldn't carry out language identification experiments, by lack of time.

- Optimization of HMM learning iteration (4.2.1)
- Optimization of fuzziness and size of templates (5)
- Evaluation of tree-based target generation technique (5)
- New target value functions (eg. Gaussian function)
- New global optimization criteria (ML, MCE)
- Language recognition application

ACKNOWLEDGMENTS

I would like to thank ATR Interpreting Telephony Laboratories for enabling me to achieve this research, and especially Dr. Akira Kurematsu for welcoming me in these laboratories, Shigeki Sagayama for welcoming me in the Speech Processing Department, Dr. Masahide Sugiyama for his very valuable explanations and advises at every moment, Yoshinaga Kato for his continuous explanations and support. I would also like to thank all the members of the department for full support.

Bibliography

[1] David P. Morgan, Christopher L. Scofield, Neural Networks and Speech Processing,Kluwer Academic Publishers, 1991.

[2] J.Potage, Reconnaissance de la parole, 1986, Polycopie ENST.

[3] X.D.Xuang, Y.Ariki, M.A. Jack, Hidden Markov Models for Speech Recognition, Edinburgh University Press, 1990.

[4] Thomas Parsons, Voice and Speech Processing, McGraw-Hill Book Company, 1987.

[5] S.J.Young,*HTK : Hidden Markov Model Toolkit V1.2,Installation guide*, Cambridge University Engineering Department, Speech Group, Dec. 1990.

[6] L.R.Rabiner, B.H.Juang, An Introduction to Hidden Markov Models, IEEE ASSP Magazine (Jan. 1986).

[7] K.Fukuzawa, Y.Kato and M.Sugiyama, "A Fuzzy Partition Model (FPM) neural network architecture for speaker-independent continuous speech recognition," Int. Conf. Spoken Lang. Processing, Banff, Oct. 1992.

[8] T.Kobayashi, Y.Uchiyama, J.Osada and K.Shirai, "Speaker adaptive phoneme recognition based on feature mapping from spectral domain to probabilistic domain," Proc. IEEE Int. Conf. Acoust, Speech, Signal Processing, San Francisco, I, pp.457–460, Mar. 1992.

[9] T.Kawabata, "Predictor codebooks speaker-independent speech recognition," Acoust. Soc. Japan, Fall Meetings, 2-P-14, pp.165–166, Oct. 1991 (in Japanese).

[10] T.Kawahara, T.Ogawa, S.Kitazawa and S.Doshita, "Phoneme recognition by combining Bayesian linear discriminations of selected pairs of classes," Proc. IEEE Int. Conf. Acoust, Speech, Signal Processing, 7.9, 1990.

[11] T.Nitta and S.Tanaka,"A comparison of subword discrimination method in speaker independent continuous speech recognition, " IEICE Technical Report, SP92, 22, pp.69–76, June, 1992 (in Japanese).

[12] Y.Tan and T.Ejima, "A network with multipartitioning units," Proc. IEEE/INNS Int. Joint Conf. Neural Networks, Washington, D.C., II, pp.439–442, June 1989.

[13] Y.Kato and M.Sugiyama, "Fuzzy Partition Models and their effect in continuous speech recognition," Proc. IEEE Workshop on Neural Networks for Signal Processing, Aug. 1992.

[14] Y.Kato and M.Sugiyama, "Speaker-independent Features Extracted by a Neural Network",*to appear in* ICASSP 93.

[15] Y.Kato, M.Sugiyama, Fuzzy Partition Models and Their Incremental Training for Continuous Speech Recognition, ASJ(E), Vol.13, No.6, pp.411-418 (1992 Nov).

[16] M.Sugiyama, Language Recognition Using Spectral Features, Proc. of Speech Research Symposium (June 1992).

[17] H.Iwamida, S.Katagiri, E.McDermott, Speaker-independent Phoneme Recognition using an LVQ/HMM hybrid, ASJ Fall meeting, 1-8-18, pp.35-36 (Sep. 1990) (in Japanese).

Appendix A

Abreviations

Table A.1: Main abreviations

FPM	Fuzzy Partition Model
AB	Alignment-based neural network : training by the "automatic alignment" procedure described
LB CI, CD	Label-based : conventional phoneme classifier ([7]) Category Independent, Dependent : classification by target signal generation, dependent or independent of phoneme categories

Here are the main notations. In this report, the input vectors are noted u and the outputs a. D stands for Kullback divergence and d for Euclidean distance.

Table A.2: Notations

S, A	respectively relative to standard and arbitrary speakers
F	fuzzy memebership function (target calculation
f	function of the network
U^S	set of speech segments for the standard speaker
U^A	set of speech segments for any arbitrary speaker

Appendix B

The FPM neural network

The FPM is a multi layer, feed forward perceptron type neural network. The FPM has multiple input/output units, which can each have an arbitrary number of dimensions. ([12]).

The particularities of a NN are the structure of a cell and the architecture of the network. The cell is the basic element to be described.

B.1 The FPM unit or cell

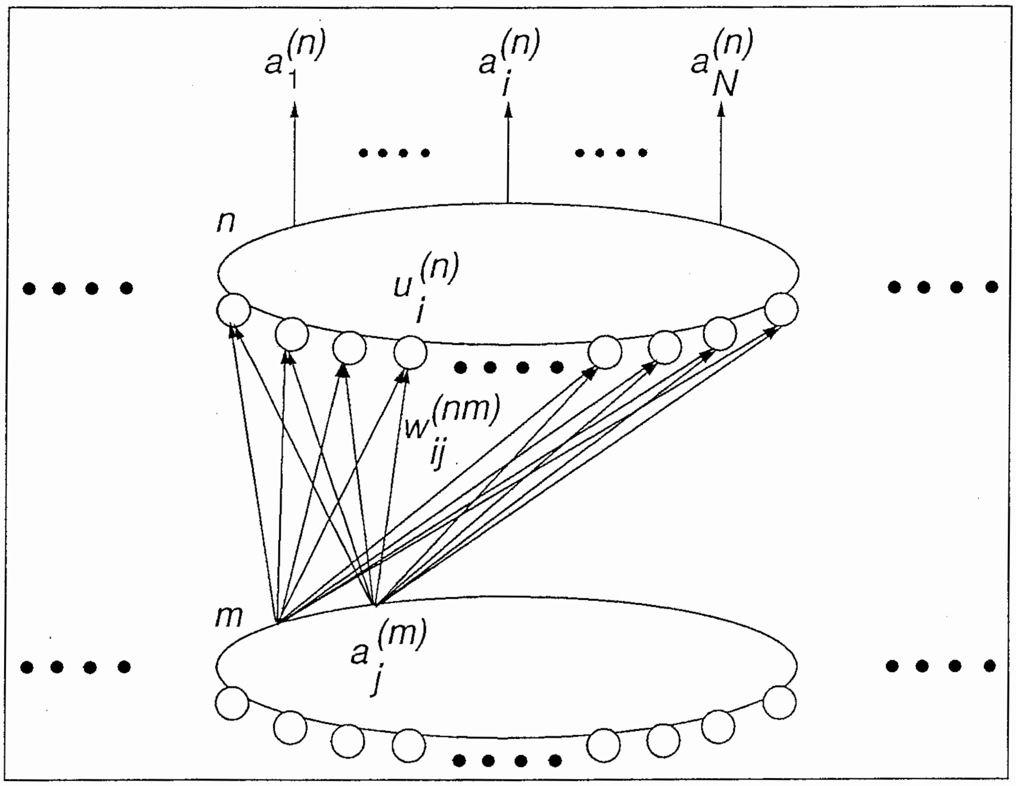


Figure B.1: An N-dimensional FPM unit

It's unit or cell is the particularity of the FPM, as compared with the classical perceptron model. Each unit can have any number of outputs N - **N-degree** unit. So each unit classifies the inputs into N categories. The particular constraint for the FPM unit is that the outputs has to be positive and their sum is constrained to 1.

According to Fig. B.1, we consider $a_i^{(n)}$ the ith output of the unit. The constraints described bellow are given by :

$$a_i^{(n)} \geq 0 \;\; \forall i \in [1, .., N] \tag{B.1}$$

$$\sum_{i=1}^{N} a_i^{(n)} = 1 \tag{B.2}$$

Given these constraints, the liberty degree of a *N-degree* unit is $N-1$: it makes $N-1$ weighted sums of inputs, then a N-dimensional output vector.

So, if $U = \{u_1, ..., u_{N-1}\}$ weighted sum of inputs, the outputs are calculated as follows :

$$a_i^{(n)} = \frac{\exp(u_i)}{1 + \sum_{k=1}^{N-1} \exp(u_k)}, \;\; \forall i \in [1, .., N-1] \tag{B.3}$$

$$a_N^{(n)} = \frac{1}{1 + \sum_{k=1}^{N-1} \exp(u_k)} \tag{B.4}$$

The inputs are simply calculated by :

$$u_k^{(n)} = \sum_j \sum_p w_{kj}^{np} a_j^{(p)} \tag{B.5}$$

where $a_j^{(p)}$ an output of the previous layer, and w_{kj}^{np} is the weight connecting the ith input of the current unit (n) and the jth of the pth unit of the "lower", adjacent layer.

So, each unit can have it's particular number of outputs. In order to simplify, each unit of a layer has the same degree.

B.2 Training

The second particularity of the FPM is the use of the Kullback divergence as error function in the learning algorithm. The learning is based on the changing of the weights with the gradient descendent (see Back Propagation Model). It was shown ([12]) that the use of this divergence increases the learning speed. The learning algorithm is a back-propagation one. The weight correction is given by :

$$\Delta w_{ij}^{nm}(t) = \eta \delta_i^{(n)} a_i^{(m)} + \alpha \Delta w_{ij}^{(nm)}(t-1) \tag{B.6}$$

. η is called the learning rate, and α the momentum rate. $\delta_i^{(n)}$ is quantitizing the difference between the layer considered and the previous one. It's expression differs for each layer. There is no theoretical mean to find the best values for α and η. In our study, α was fixed to 0.9([12]) and η to 0.001.

Appendix C

Kullback divergence

The Kullback theory is applied to probabilistic distributions. So, the application of the Kullback divergence as a distance measure requires some constraints. In the FPM, the given constraint being that the sum of outputs has to be equal to 1, and all the outputs has to be positive. So, the outputs can be considered as a probability distribution ([12]). The error measure can be interpreted as the difference between the desired stochastic distribution and the accomplished one.

The mathematical formula is given by:

$$D(T^{(n)}, A^{(n)}) = \sum_{i=1}^{N} t_i^{(n)} \log \frac{t_i^{(n)}}{a_i^{(m)}} \tag{C.1}$$

measuring the distance between the target vector (which is the desired value) $T^{(n)}$and the actual output of the Neural network $A^{(n)}$/.

Appendix D

Recognition results per phoneme

Table D.1: Amount of test data, MAU and MNM, 216 balanced words per speaker

Phoneme	Nb of phonemes/speaker		All
	MAU	MNM	MAU+MNM
p1	6	6	12
p2	7	7	14
t1	10	10	20
t2	19	19	38
k1	21	21	42
k2	51	50	101
b1	7	7	14
b2	20	19	39
d1	9	9	18
d2	12	11	23
g1	8	8	16
ng	17	18	35
m	35	35	70
n	26	26	52
N	45	38	83
r	61	61	122
w	12	10	22
y	14	15	29
s	31	31	62
sh	9	9	18
h	21	21	42
z	26	23	49
ch1	1	1	2
ch2	9	9	18
ts1	2	2	4
ts2	6	6	12
sy	14	13	27
hy	11	11	22
zy	14	13	27
cy	11	11	22
py	7	7	14
ky	11	11	22
by	10	9	19
gy1	4	4	8
ngy	5	5	10
my	9	9	18
ny	9	9	18
ry	11	11	22
aa	8	8	16
ii	9	9	18
uu	38	38	76
eei	14	15	29
oou	55	55	110
a	161	159	320
i	100	100	200
u	116	112	228
e	78	78	156
o	107	107	214
sil	432	432	864

Table D.2: Test data without "unknown labels" files

Phoneme	Nb of phonemes per speaker		All
	MAU	MNM	MAU+MNM
p1	5	5	10
p2	7	7	14
t1	9	9	18
t2	19	19	38
k1	19	16	35
k2	41	38	79
b1	4	4	8
b2	20	19	39
d1	8	8	16
d2	12	11	23
g1	7	7	14
ng	16	16	32
m	32	28	60
n	26	25	51
N	44	37	81
r	53	50	103
w	12	9	21
y	13	14	27
s	27	25	52
sh	9	8	17
h	18	18	36
z	25	21	46
ch1	1	1	2
ch2	8	9	17
ts1	1	1	2
ts2	4	5	9
sy	13	12	25
hy	11	10	21
zy	11	11	22
cy	11	11	22
py	7	7	14
ky	11	11	22
by	10	9	19
gy1	4	4	8
ngy	5	5	10
my	6	6	12
ny	9	9	18
ry	10	10	20
aa	7	6	13
ii	9	8	17
uu	35	37	72
eei	12	13	25
oou	54	52	106
a	140	127	267
i	89	87	177
u	92	85	177
e	70	67	137
o	93	90	183
sil	376	360	736

Table D.3: Recognition per phoneme (g=1.6, M=4)

Phoneme	Correct	Total	Recog. rate
a	1440	1778	80.98
i	904	1312	69.90
u	1080	1511	71.47
e	416	675	62.63
o	698	914	76.37
p1	1	14	7.14
p2	0	28	0
t1	153	201	76.12
t2	185	238	77.73
k1	268	447	59.95
k2	444	774	57.36
b1	39	60	65.00
b2	124	157	78.98
d1	62	73	84.93
d2	107	124	86.29
g1	36	71	50.70
ng	146	198	73.73
m	246	493	49.89
n	140	271	51.66
N	423	472	89.61
r	563	785	71.71
w	75	79	94.93
y	139	164	84.75
s	380	474	81.17
sh	224	256	87.5
h	221	318	69.50
z	148	191	77.49
ch1	8	22	36.36
ch2	64	77	69.64
ts1	39	56	89.52
ts2	188	210	62.50
sy	80	128	81.25
hy	13	16	81.25
zy	59	83	71.09
cy	49	62	79.03
py	0	3	0
ky	49	54	79.63
by	3	9	33.33
gy1	4	7	57.14
ngy	2	9	22.22
my	4	9	44.44
ny	6	9	66.66
ry	37	47	78.72
aa	18	18	100
ii	59	60	98.33
uu	106	150	92.58
eei	155	169	91.71
oou	474	512	93.54
sil	5038	5386	93.22

Table D.4: Recognition per phoneme (g=1.6, M=6)

Phoneme	Correct	Total	Recog. rate
a	1471	1778	82.73
i	960	1312	73.17
u	1076	1511	71.21
e	422	675	62.51
o	705	914	77.13
p1	5	14	35.71
p2	12	28	42.86
t1	160	201	79.60
t2	210	238	88.23
k1	318	447	71.14
k2	542	774	70.02
b1	47	60	78.33
b2	126	157	80.25
d1	58	73	79.45
d2	105	124	84.67
g1	44	71	61.97
ng	151	198	76.26
m	247	493	50.10
n	166	271	61.25
N	428	472	90.67
r	579	785	73.76
w	77	79	97.45
y	142	164	86.58
s	388	474	81.85
sh	221	256	86.32
h	195	318	61.32
z	155	191	81.15
ch1	1	22	4.54
ch2	68	77	88.31
ts1	30	56	53.57
ts2	195	210	92.85
sy	91	128	71.09
hy	13	16	81.25
zy	68	83	81.92
cy	46	62	74.19
py	1	3	33.33
ky	42	54	77.78
by	3	9	33.33
gy1	3	7	42.86
ngy	3	9	33.33
my	4	9	44.44
ny	6	9	66.66
ry	38	47	80.85
aa	18	18	100
ii	60	60	100
uu	112	150	74.66
eei	155	169	91.71
oou	478	512	93.35
sil	5021	5386	93.22

Table D.5: Recognition per phoneme (g=1.6, M=8)

Phoneme	Correct	Total	Recog. rate
a	1466	1703	86.08
i	921	1270	72.51
u	1050	1465	71.67
e	398	640	62.18
o	685	863	79.37
p1	1	8	12.5
p2	0	22	0.00
t1	155	196	79.08
t2	202	229	88.21
k1	302	438	68.95
k2	518	750	60.06
b1	46	59	79.96
b2	125	148	84.46
d1	55	68	80.88
d2	98	117	83.76
g1	47	68	69.11
ng	153	191	80.10
m	276	476	57.98
n	163	260	62.69
N	415	456	91.00
r	563	757	74.37
w	66	73	90.41
y	136	159	85.54
s	384	461	83.29
sh	224	250	89.59
h	222	309	71.84
z	151	182	82.97
ch1	1	21	4.76
ch2	67	72	93.05
ts1	17	55	30.90
ts2	193	207	93.24
sy	91	122	74.59
hy	9	9	100
zy	63	75	84.00
cy	43	55	78.18
py	0	0	
ky	41	48	85.41
by	1	4	25
gy1	2	5	40.00
ngy	3	7	42.86
my	2	4	50.00
ny	3	4	75.00
ry	35	40	87.50
aa	12	12	100
ii	53	54	98.14
uu	106	131	80.91
eei	147	159	92.45
oou	456	483	94.40
sil	4819	5178	93.06

Table D.6: Recognition per phoneme (g=1.6, M=12)

Phoneme	Correct	Total	Recog. rate
a	1553	1778	87.34
i	999	1312	76.14
u	1104	1511	73.06
e	448	675	66.37
o	738	914	80.74
p1	0	14	0
p2	0	28	0
t1	158	201	78.60
t2	219	238	92.01
k1	307	447	68.68
k2	617	774	79.71
b1	47	60	78.33
b2	133	157	84.71
d1	62	73	84.93
d2	111	124	89.51
g1	47	71	66.19
ng	154	198	77.79
m	312	493	63.28
n	162	271	59.77
N	430	472	91.10
r	576	785	73.37
w	77	79	97.47
y	147	164	89.63
s	409	474	86.28
sh	228	256	89.06
h	239	318	75.15
z	156	191	81.67
ch1	4	22	18.18
ch2	67	77	87.01
ts1	21	56	37.5
ts2	199	210	94.76
sy	105	128	82.03
hy	12	16	75.00
zy	73	83	87.95
cy	49	62	79.03
py	1	3	33.33
ky	47	54	87.04
by	3	9	33.33
gy1	2	7	28.57
ngy	3	9	33.33
my	4	9	44.44
ny	6	9	66.66
ry	40	47	85.10
aa	18	18	100
ii	59	60	98.33
uu	107	150	71.36
eei	154	169	91.12
oou	480	512	93.75
sil	5026	5386	93.32

Table D.7: Recognition per phoneme (g=1.6, M=4, templates=10)

Phoneme	Correct	Total	Recog. rate
a	1481	1778	83.29
i	943	1312	71.87
u	1032	1511	68.29
e	445	675	65.92
o	712	914	77.89
p1	0	14	0
p2	0	28	0
t1	159	201	79.10
t2	210	238	88.23
k1	262	447	58.61
k2	519	774	67.05
b1	46	60	76.66
b2	120	157	76.43
d1	62	73	84.93
d2	103	124	83.06
g1	42	71	59.15
ng	155	198	78.28
m	226	493	45.84
n	150	271	55.35
N	415	472	87.92
r	563	785	71.72
w	70	79	88.60
y	139	164	84.75
s	393	474	82.91
sh	231	256	90.23
h	200	318	62.89
z	143	191	74.87
ch1	9	22	40.90
ch2	57	77	74.02
ts1	34	56	60.71
ts2	188	210	89.52
sy	63	128	49.21
hy	13	16	81.25
zy	68	83	81.92
cy	46	62	74.19
py	0	3	0.00
ky	45	54	77.78
by	4	9	44.44
gy1	4	7	57.14
ngy	3	9	33.33
my	5	9	55.55
ny	6	9	66.66
ry	37	47	78.72
aa	18	18	100
ii	59	60	98.33
uu	105	150	70.00
eei	152	169	89.94
oou	490	512	95.70
sil	5032	5386	93.43

Appendix E

Confusion Matrix for Phoneme Recognitions

1. CI, $M = 25$, mixture: 4, $g = 1.1$, open speakers
2. CI, $M = 25$, mixture: 6, $g = 1.1$, open speakers
3. CI, $M = 25$, mixture: 8, $g = 1.1$, open speakers
4. CI, $M = 25$, mixture: 12, $g = 1.1$, open speakers
5. CI, $M = 25$, mixture: 16, $g = 1.1$, open speakers
6. CI, $M = 25$, mixture: 4, $g = 1.1$, standard speaker
7. CI, $M = 25$, mixture: 8, $g = 1.1$, standard speaker
8. CI, $M = 25$, mixture: 12, $g = 1.1$, standard speaker
9. CI, $M = 25$, mixture: 4, $g = 1.6$, standard speaker
10. CI, $M = 25$, mixture: 6, $g = 1.6$, standard speaker
11. CI, $M = 25$, mixture: 8, $g = 1.6$, standard speaker
12. CI, $M = 10, L = 5$, mixture: 4, $g = 1.6$, standard speaker

CI, 25 templates
TEST DATA : 368 words : - MAU 188
- MNM 180
Number of mixtures : 4

------------------------ Overall Results ------------------------
PHRASE: %Correct=0.54 [H=2, S=366, N=368]
PHONE: %Corr=66.76, Acc=50.27 [H=1984, D=86, S=902, I=490, N=2972]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	196	0	1	0	9	0	0	0	1	0	0	0	0	0	0	0	0	1	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	0	0	0	1	4	[71 errs]
i	0	82	2	1	0	0	0	1	0	0	0	0	0	0	0	1	0	3	1	2	0	0	6	0	0	0	1	0	0	0	0	0	0	1	0	0	2	0	0	0	0	11	1	0	41	1	2	1	8	[86 errs]
u	0	2	105	1	1	1	0	0	1	0	0	0	0	0	0	1	3	7	0	16	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	3	1	2	0	0	2	5	0	2	6	[59 errs]
e	0	3	23	47	2	0	0	0	1	0	0	0	2	1	0	2	5	7	0	3	0	1	2	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	4	0	1	0	11	12	1	4	[88 errs]
o	1	0	5	0	77	0	0	0	0	0	0	0	1	0	0	0	2	3	0	4	1	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	68	10	[101 errs]
p1	0	0	0	0	0	3	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
p2	0	0	0	0	0	1	3	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[10 errs]
t1	0	0	0	0	0	0	0	9	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
t2	0	0	0	0	0	0	1	0	34	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[3 errs]
k1	0	0	1	0	0	6	0	6	1	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	[19 errs]
k2	0	0	0	0	1	3	3	2	9	7	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	12	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	1	1	0	1	[50 errs]
b1	0	0	0	0	0	0	0	0	3	0	0	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
b2	0	0	0	0	0	0	0	0	0	0	0	3	25	0	4	0	0	1	2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[12 errs]
d1	0	0	0	0	0	0	0	0	2	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[3 errs]
d2	0	0	1	0	0	0	0	0	0	0	0	1	5	0	11	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	[12 errs]
g1	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	[7 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	11	8	2	2	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0	0	[20 errs]
m	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	3	1	25	5	4	2	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	1	0	0	0	1	0	[29 errs]
n	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	1	5	11	13	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0	4	0	[32 errs]
N	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	2	2	0	66	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	2	[14 errs]
r	0	0	1	0	0	0	0	1	1	0	0	1	5	2	0	2	8	9	4	3	30	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	3	4	1	1	2	0	3	1	[55 errs]
w	1	0	0	0	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
y	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	1	14	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	1	0	0	0	1	0	0	0	[13 errs]
s	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	36	2	0	1	0	1	4	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[15 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
h	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	1	0	0	0	0	1	1	1	20	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[14 errs]
z	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	33	0	1	0	0	0	0	2	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	[11 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[1 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	12	0	0	4	0	4	0	2	0	0	0	0	0	0	0	0	1	0	[13 errs]
hy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	6	0	4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	[15 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	17	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	3	0	0	11	0	5	0	0	1	0	0	0	0	0	0	0	0	0	[11 errs]
py	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	3	4	0	1	0	0	0	0	0	0	0	0	0	0	[11 errs]
ky	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	1	0	11	0	1	0	0	0	0	0	0	0	0	0	0	[10 errs]
by	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	7	1	2	2	1	1	0	0	0	0	0	0	[12 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	7	0	0	0	0	0	0	0	0	0	0	[1 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	1	1	0	0	0	0	0	0	[2 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5	0	0	1	0	0	0	0	0	[7 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	[0 errs]
ry	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	2	4	0	1	0	0	0	0	[16 errs]
aa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	[2 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	[0 errs]
uu	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	0	0	2	[18 errs]
eei	0	2	3	4	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	10	0	1	[15 errs]
oou	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	105	0	[1 errs]
sil	0	0	0	0	0	0	1	0	10	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	719	[13 errs]

... done
End of Stage 6

CI, 25 templates
TEST DATA : 368 words : - MAU 188
- MNM 180
Number of mixtures : 6

------------------------ Overall Results ------------------------
PHRASE: %Correct=1.09 [H=4, S=364, N=368]
PHONE: %Corr=68.44, Acc=45.05 [H=2034, D=65, S=873, I=695, N=2972]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	195	0	2	1	6	0	1	0	2	0	0	0	2	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	39	0	1	0	1	9	[69 errs]
i	0	89	1	0	0	0	0	0	0	2	0	0	0	0	0	0	3	2	1	2	1	1	9	0	0	0	0	0	0	0	0	0	1	1	0	0	2	0	0	1	0	7	0	0	21	2	0	1	21	[79 errs]
u	0	1	103	0	3	2	0	0	0	0	0	0	0	0	0	1	3	2	1	13	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	1	0	0	0	1	0	3	0	0	3	9	1	2	15	[64 errs]
e	1	3	21	55	3	0	0	0	2	0	0	0	0	0	0	1	2	5	2	5	1	0	2	0	0	0	2	0	0	0	0	0	0	0	0	1	0	0	1	0	0	2	0	2	0	13	4	1	6	[80 errs]
o	0	0	5	0	81	0	1	1	2	0	1	0	3	0	0	0	2	2	0	3	1	5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64	10	[101 errs]
p1	0	0	0	0	0	1	0	3	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	[9 errs]
p2	0	0	0	0	0	1	3	0	7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	[11 errs]
t1	0	0	0	0	1	2	0	5	1	3	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	[12 errs]
t2	0	0	0	0	0	0	1	0	35	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[3 errs]
k1	0	0	0	0	0	1	0	3	1	22	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	[12 errs]
k2	0	0	0	1	0	0	5	1	5	5	30	0	0	0	0	0	1	0	1	0	0	0	1	0	0	4	0	0	9	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	4	0	1	3	1	[46 errs]
b1	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	1	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[6 errs]
b2	0	0	0	0	1	0	0	0	0	0	0	0	26	1	5	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	[11 errs]
d1	0	0	0	0	0	0	0	0	2	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	[3 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	4	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	[7 errs]
g1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	5	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	1	0	0	1	0	0	[9 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	10	4	3	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	3	2	0	0	[18 errs]
m	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	6	22	12	2	0	2	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	2	0	1	0	0	1	3	2	[36 errs]
n	0	0	3	0	0	0	0	0	1	0	0	0	2	0	0	3	2	5	18	1	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	2	0	1	2	[30 errs]
N	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	3	2	1	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	2	[15 errs]
r	0	0	7	2	2	0	0	0	0	0	0	0	4	2	3	5	8	2	6	2	29	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	2	2	0	2	1	0	3	1	[58 errs]
w	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	[5 errs]
y	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5	1	0	0	1	1	0	0	0	1	[13 errs]
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	36	0	1	0	0	0	4	2	0	2	0	2	0	0	0	0	0	0	1	0	0	0	1	0	0	2	[16 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	[1 errs]
h	0	0	0	0	1	1	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	23	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	[11 errs]
z	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	37	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	[7 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[1 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	15	2	0	4	0	0	0	1	0	0	0	0	0	0	0	0	1	0	[10 errs]
hy	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	9	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	[12 errs]
zy	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	18	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	1	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
py	0	2	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	3	2	0	0	0	0	0	0	0	0	0	0	0	0	[11 errs]
ky	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	3	0	14	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
by	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	1	1	2	2	0	0	0	2	0	0	0	[12 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	[2 errs]
ngy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	2	0	0	0	1	0	1	0	[5 errs]
my	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	4	1	0	1	0	0	0	0	0	[8 errs]
ny	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	15	0	0	0	0	0	0	0	[3 errs]
ry	1	0	2	0	1	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	1	1	5	0	0	0	0	0	0	[14 errs]
aa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	[2 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	[0 errs]
uu	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	10	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	54	0	0	2	[18 errs]
eei	0	1	0	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	4	9	0	2	[16 errs]
oou	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	103	0	[3 errs]
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	735	[0 errs]

... done
End of Stage 6

Jan 28 16:59 1993 /NFS/atrp10/p10/users/xdonescu/HTK/t25.40.8m Page 1

CI, 25 templates
TEST DATA : 368 words : - MAU 188
- MNM 180
Number of mixtures : 8

------------------------ Overall Results ------------------------
PHRASE: %Correct=1.90 [H=7, S=361, N=368]
PHONE: %Corr=70.29, Acc=48.62 [H=2089, D=72, S=811, I=644, N=2972]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	199	0	0	0	11	0	0	0	1	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	0	0	1	1	6	[67 errs]
i	0	96	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	0	1	5	0	0	0	2	0	1	0	0	0	0	1	1	0	3	1	0	1	0	7	2	0	26	1	0	0	14	[74 errs]
u	0	1	109	1	1	1	0	0	0	0	0	0	0	0	0	0	1	2	1	17	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1	2	6	0	3	12	[55 errs]
e	0	5	31	55	1	0	0	0	2	0	0	0	1	0	0	0	1	6	1	2	1	0	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	2	1	1	1	0	0	0	0	5	7	2	5	[79 errs]
o	0	0	2	0	90	0	1	0	2	0	0	0	2	0	0	0	2	1	0	4	0	2	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	57	12	[90 errs]
p1	0	0	0	0	0	1	0	1	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[8 errs]
p2	1	0	0	0	0	0	3	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[10 errs]
t1	0	0	0	0	0	1	0	5	4	2	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	[11 errs]
t2	0	0	0	0	0	0	1	0	34	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
k1	0	0	0	0	0	1	0	5	2	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[9 errs]
k2	0	0	0	0	0	1	6	3	3	5	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	9	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	3	1	0	2	1	[40 errs]
b1	0	0	0	0	0	0	1	0	2	0	0	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	[6 errs]
b2	0	0	0	0	1	0	0	0	0	0	0	0	25	1	4	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	[12 errs]
d1	0	0	0	0	0	0	0	0	2	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	[3 errs]
d2	0	0	0	0	1	0	0	0	0	0	0	0	3	1	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	[8 errs]
g1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	5	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	[9 errs]
ng	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	10	5	5	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1	0	0	1	3	0	0	0	[22 errs]
m	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	5	23	9	7	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	2	1	[33 errs]
n	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	2	5	5	21	1	1	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	6	1	[29 errs]
N	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1	69	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	[11 errs]
r	1	0	8	1	0	0	0	0	0	1	0	1	4	3	2	2	6	7	7	2	30	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	1	1	1	1	0	2	2	[57 errs]
w	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
y	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	1	1	0	0	0	0	[12 errs]
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	0	0	0	0	0	2	1	2	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	[9 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
h	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	24	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[10 errs]
z	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	38	0	0	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	[7 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[1 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
hy	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	10	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	4	1	0	15	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
py	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	4	3	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
ky	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	1	0	4	0	11	0	0	0	0	0	0	0	0	0	0	0	0	[11 errs]
by	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	2	0	3	1	2	0	0	1	0	0	0	[11 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	[2 errs]
ngy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	4	0	0	0	0	0	0	0	[6 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	2	0	1	0	0	0	0	0	[6 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	15	0	0	0	0	0	0	0	[3 errs]
ry	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1	0	1	1	7	0	0	0	0	0	0	[13 errs]
aa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	[3 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	[0 errs]
uu	0	0	7	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	5	[23 errs]
eei	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	2	14	0	2	[11 errs]
oou	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	103	0	[2 errs]
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	735	[0 errs]

... done
End of Stage 6

50

Jan 28 16:58 1993 /NFS/atrp10/p10/users/xdonescu/HTK/t25.40.12m Page 1

CI, 25 templates
TEST DATA : 368 words : - MAU 188
- MNM 180
Number of mixtures : 12

------------------------ Overall Results ------------------------
PHRASE: %Correct=0.82 [H=3, S=365, N=368]
PHONE: %Corr=71.67, Acc=50.27 [H=2130, D=66, S=776, I=636, N=2972]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	218	0	0	0	7	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	1	0	1	10	[46 errs]
i	0	110	1	0	1	0	0	0	0	0	0	0	0	0	0	0	2	1	0	1	0	1	7	1	0	0	1	0	0	0	0	0	0	1	0	0	3	0	0	1	0	6	0	0	23	1	0	0	13	[64 errs]
u	0	1	104	2	3	0	2	0	1	0	0	0	0	0	0	1	3	4	0	15	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	4	3	1	3	10	[58 errs]
e	0	6	27	45	2	0	0	0	1	0	0	0	0	0	0	1	0	1	2	6	0	1	2	0	0	0	1	0	0	0	0	0	0	2	0	0	0	1	0	1	0	3	0	0	0	10	15	1	6	[89 errs]
o	0	0	4	0	103	0	0	0	2	0	0	0	0	0	0	1	1	0	1	3	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	44	13	[76 errs]
p1	0	0	0	0	0	1	0	0	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
p2	1	0	0	0	1	0	3	0	7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	[11 errs]
t1	0	0	0	0	0	0	0	4	8	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[13 errs]
t2	0	0	0	0	0	0	1	0	34	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	[4 errs]
k1	0	0	0	0	0	0	0	3	2	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
k2	0	0	0	1	0	0	0	2	6	3	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	7	0	0	0	0	0	1	1	3	0	0	0	0	1	0	0	1	1	1	1	1	[34 errs]
b1	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	[5 errs]
b2	0	0	0	0	1	0	0	0	0	0	0	0	26	1	3	1	1	0	3	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[12 errs]
d1	0	0	0	0	0	0	0	0	3	0	0	2	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	3	0	16	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	[7 errs]
g1	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2	0	0	0	0	0	0	0	0	0	1	[8 errs]
ng	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	10	5	5	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	1	0	0	[19 errs]
m	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	4	24	7	5	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	3	0	0	0	0	3	2	1	[33 errs]
n	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	11	20	2	1	2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	4	0	[28 errs]
N	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	2	1	1	66	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	2	0	2	0	[13 errs]
r	0	0	5	0	1	0	0	0	0	0	0	0	2	1	1	3	5	9	6	1	37	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	2	2	2	2	2	4	[51 errs]
w	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
y	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	2	1	0	0	1	1	0	0	2	[18 errs]
s	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[7 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
h	0	0	0	0	0	1	1	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	[11 errs]
z	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	33	0	0	0	0	0	0	4	1	0	1	0	0	0	0	0	0	0	2	1	0	0	1	[12 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	[3 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	[2 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[6 errs]
hy	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	11	0	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	18	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
py	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	4	3	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
ky	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	14	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
by	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	8	1	1	4	1	0	0	0	0	0	0	0	[11 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	5	0	0	0	1	0	0	0	0	1	0	[3 errs]
ngy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	2	0	0	0	1	0	0	0	[4 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5	1	0	0	0	0	0	1	0	[7 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	15	0	0	0	0	0	0	0	[3 errs]
ry	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	0	1	0	6	0	0	1	0	0	1	[14 errs]
aa	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	[4 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	[0 errs]
uu	0	1	2	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51	0	0	3	[19 errs]
eei	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	5	12	0	2	[13 errs]
oou	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	104	0	[2 errs]
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	735	[0 errs]

... done
End of Stage 6

CI, 25 templates
TEST DATA : 368 words : - MAU 188
- MNM 180
Number of mixtures : 16

------------------------ Overall Results ------------------------
PHRASE: %Correct=2.99 [H=11, S=357, N=368]
PHONE: %Corr=71.47, Acc=55.85 [H=2124, D=91, S=757, I=464, N=2972]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	223	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	29	0	0	0	2	0	[41 errs]
i	0	107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	2	1	1	7	1	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	2	0	4	0	0	29	1	0	0	7	[62 errs]
u	0	2	106	0	1	1	0	0	0	0	0	0	0	0	0	1	1	3	1	14	0	0	0	1	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	3	14	1	4	5	[56 errs]
e	0	5	26	34	2	0	0	0	2	0	0	0	0	0	0	1	2	4	0	3	1	0	1	0	0	0	0	0	0	0	0	0	1	2	0	0	1	0	0	0	1	1	0	1	0	10	26	2	5	[97 errs]
o	0	0	2	0	102	0	0	0	1	0	0	0	0	0	0	1	1	1	2	3	2	4	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	5	[75 errs]
p1	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[9 errs]
p2	0	0	0	0	0	0	3	0	8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[11 errs]
t1	0	0	1	0	0	0	0	3	3	4	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[13 errs]
t2	0	0	0	0	0	0	0	0	34	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
k1	0	0	0	0	0	0	1	3	0	24	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
k2	0	0	0	0	0	0	1	2	6	3	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	9	0	0	0	0	0	2	2	3	0	0	0	0	0	0	0	0	1	0	0	1	[36 errs]
b1	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	[6 errs]
b2	0	0	0	0	0	0	0	0	0	0	1	0	27	0	3	1	0	0	4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	[11 errs]
d1	0	0	0	0	0	0	0	0	1	0	0	1	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	[3 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	3	0	15	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	[8 errs]
g1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	7	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	5	4	0	1	1	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	3	0	0	0	[20 errs]
m	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	4	32	6	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	1	1	0	[22 errs]
n	0	0	3	0	0	0	0	0	1	0	0	0	2	0	0	0	2	6	23	0	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	4	0	[24 errs]
N	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	5	2	0	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	4	0	[14 errs]
r	0	0	4	0	1	0	0	0	1	0	1	1	4	2	1	2	3	7	7	1	37	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	1	0	1	4	0	[47 errs]
w	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	[6 errs]
y	0	3	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	9	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	3	1	0	0	0	0	0	0	0	0	[16 errs]
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	44	1	0	1	0	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
h	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	26	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[9 errs]
z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	38	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	[7 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	2	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	[2 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[3 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	19	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	[5 errs]
hy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	12	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[9 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	20	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4	1	0	13	0	2	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
py	0	2	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	0	0	0	0	0	0	0	0	0	0	0	1	[12 errs]
ky	0	3	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	11	0	0	0	0	0	0	0	0	0	0	0	0	[11 errs]
by	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	9	0	0	2	1	0	0	0	1	0	0	0	[10 errs]
gy1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	3	0	0	0	1	0	0	0	0	0	0	[5 errs]
ngy	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	3	0	0	0	0	0	0	0	[7 errs]
my	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5	1	0	0	0	0	0	0	0	[7 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	16	0	0	0	0	0	0	0	[2 errs]
ry	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	1	2	5	0	0	0	0	0	0	[15 errs]
aa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	[2 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	[0 errs]
uu	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	10	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	0	0	1	[16 errs]
eei	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	14	0	2	[11 errs]
oou	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	105	0	[1 errs]
sil	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	733	[0 errs]

... done
End of Stage 6

52

STANDARD SPEAKER
TEST DATA
FUZZINESS 1.1
TEMPLATES 25
MIXTURES : 4

------------------------ Overall Results ------------------------
PHRASE: %Correct=5.54 [H=114, S=1943, N=2057]
PHONE: %Corr=74.49, Acc=58.97 [H=11066, D=174, S=3616, I=2306, N=14856]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	1156	0	0	7	16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	16	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	133	0	1	1	1	6	[188 errs]
i	0	470	1	32	0	0	8	0	2	2	0	0	0	0	0	0	5	0	0	51	0	1	25	1	3	3	2	0	4	0	2	1	1	0	0	3	3	0	0	7	1	1	3	0	240	29	25	1	27	[484 errs]
u	0	6	735	17	8	0	2	1	3	0	0	0	1	0	0	0	8	13	1	104	4	1	2	1	1	0	0	0	1	4	4	0	1	0	0	0	5	0	0	1	0	2	1	0	3	118	21	15	13	[362 errs]
e	0	1	5	284	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	4	0	0	4	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	2	0	1	8	151	0	17	[200 errs]
o	7	0	3	2	373	0	1	1	6	0	1	0	0	0	0	0	0	5	1	0	1	3	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	181	13	[228 errs]
p1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
p2	0	0	0	0	0	0	6	2	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[12 errs]
t1	0	0	0	0	0	1	5	158	11	2	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[22 errs]
t2	1	0	0	0	0	0	28	6	141	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[40 errs]
k1	0	0	0	0	1	1	1	78	9	223	6	0	0	0	0	0	0	0	0	0	0	4	0	1	1	27	0	5	3	6	0	0	0	0	3	0	9	0	0	0	0	0	0	1	0	0	1	2	1	[160 errs]
k2	1	0	1	2	0	0	15	146	83	101	157	0	0	0	0	0	1	0	0	1	0	1	0	3	1	8	0	1	4	0	11	0	0	0	2	0	22	0	0	0	0	0	0	2	7	1	2	4	13	[433 errs]
b1	0	0	0	0	0	0	0	0	1	0	0	11	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[3 errs]
b2	0	0	0	0	0	0	0	0	1	0	0	5	72	0	29	0	0	1	3	1	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	1	[50 errs]
d1	0	0	0	0	0	0	0	0	1	0	0	4	0	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[5 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	1	11	3	67	1	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	[20 errs]
g1	0	0	0	0	0	1	0	8	0	0	0	4	0	7	0	22	0	0	0	0	0	0	8	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	[32 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	119	4	5	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	3	0	0	2	3	2	0	0	[26 errs]
m	0	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	79	70	50	15	4	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	1	0	0	0	6	4	1	4	0	[174 errs]
n	0	0	1	0	0	0	0	0	12	0	0	3	2	0	0	0	33	23	103	3	4	9	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	5	6	0	0	4	1	0	1	0	[112 errs]
N	0	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	334	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	7	2	0	0	[23 errs]
r	0	5	1	1	1	0	0	0	3	0	0	0	8	0	1	0	8	3	8	5	377	23	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3	0	0	36	0	5	8	4	1	0	[127 errs]
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	28	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
y	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	0	[7 errs]
s	0	0	1	0	0	0	0	7	2	1	2	0	0	0	0	0	1	0	0	0	0	0	0	259	6	0	0	1	1	79	5	6	4	0	8	0	0	0	0	0	0	0	0	0	0	0	2	0	1	[127 errs]
sh	0	0	0	0	0	0	1	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	172	0	0	21	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	[30 errs]
h	1	0	0	0	0	0	1	1	1	9	0	0	0	0	0	0	0	0	0	1	0	0	0	1	15	116	0	13	6	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	[53 errs]
z	0	0	0	0	0	0	0	0	1	0	1	0	4	2	0	6	0	0	1	1	2	1	0	0	0	0	107	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	2	0	0	1	2	[30 errs]
ch1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[7 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	[6 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	43	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[7 errs]
ts2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	2	0	1	0	0	2	43	114	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	[55 errs]
sy	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	6	2	5	0	0	0	1	5	1	55	7	0	7	3	1	0	0	1	0	0	0	0	0	0	6	0	0	[50 errs]
hy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[1 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0	56	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	[6 errs]
cy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1	3	0	32	4	2	0	0	0	0	0	0	0	0	0	3	0	0	[16 errs]
py	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
ky	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6	1	34	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
by	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	[2 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	[3 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	0	0	0	0	0	0	0	0	[3 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	[2 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	[1 errs]
ry	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	1	0	0	0	1	[4 errs]
aa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	1	[3 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	46	0	0	0	0	[0 errs]
uu	0	0	5	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	69	7	1	5	[29 errs]
eei	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	117	0	0	[6 errs]
oou	0	0	2	0	10	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	353	3	[18 errs]
sil	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3832	[9 errs]

... done
End of Stage 6

```
STANDARD SPEAKER
TEST DATA 1702 WORDS
FUZZINESS 1.1
TEMPLATES 25
MIXTURES 8
------------------------ Overall Results ------------------------
PHRASE: %Correct=7.14 [H=147, S=1912, N=2059]
PHONE:  %Corr=79.33, Acc=63.38 [H=11797, D=175, S=2898, I=2372, N=14870]
-----------------------------------------------------------------
Confusion Matrix
       a  i  u  e  o  p  p  t  t  k  k  b  b  d  d  g  n  m  n  N  r  w  y  s  s  h  z  c  c  t  t  s  h  z  c  p  k  b  g  n  m  n  r  a  i  u  e  o  s
                      1  2  1  2  1  2  1  2  1  2  1  g                       h        h  h  s  s  y  y  y  y  y  y  y  y  y  g  y  y  y  a  i  u  e  o  i
                                                                                        1  2  1  2                          1  y                    i  u  1
   a 1237 0 0 1 9 0 0 0 0 0 0 0 0 0 0 0 0 2 1 0 0 4 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 83 1 0 0 1 7 [112 errs]
   i 0735 5 5 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0 7 0 1 24 1 1 0 0 0 0 0 0 2 0 0 1 0 9 0 0 2 1 1 1 0109 23 30 0 4 [230 errs]
   u 1 23750 15 16 0 0 0 2 0 0 0 1 0 0 0 2 13 3 94 5 0 1 0 3 0 1 0 1 6 2 1 0 0 1 0 2 0 0 0 0 1 0 0 0111 20 15 14 [354 errs]
   e 0 3 2266 0 0 0 1 2 0 0 0 0 0 0 0 1 0 1 2 0 0 2 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 6174 1 20 [218 errs]
   o 2 1 1 2407 0 0 0 5 1 2 1 0 0 0 0 0 1 2 1 1 3 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1152 5 [184 errs]
  p1 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  0 errs]
  p2 0 0 0 0 0 0 4 0 10 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 [ 14 errs]
  t1 0 0 0 0 0 0 0145 25 6 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 [ 34 errs]
  t2 0 0 0 0 0 0 8 0171 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 [ 11 errs]
  k1 1 0 0 0 1 0 2 30 4311 7 0 0 0 0 0 0 0 0 0 0 0 0 1 4 4 0 0 5 1 1 1 0 0 2 0 5 0 0 0 0 0 0 0 0 1 0 1 3 [ 74 errs]
  k2 1 0 0 2 0 0 5 96 81 85264 0 0 0 0 0 0 0 0 0 0 0 0 1 0 7 0 0 3 0 6 0 0 0 4 0 21 0 0 0 0 0 0 4 2 0 0 2 7 [327 errs]
  b1 0 0 0 0 0 0 0 0 1 0 0 12 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  2 errs]
  b2 0 0 0 0 0 0 0 0 1 0 0 3 88 1 21 0 1 1 1 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 2 0 [ 35 errs]
  d1 0 0 0 0 0 0 0 0 1 0 0 5 0 53 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  6 errs]
  d2 0 0 0 0 0 0 0 0 0 0 0 0 18 3 63 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 [ 24 errs]
  g1 0 0 0 0 0 0 0 2 0 0 0 7 0 11 0 27 0 0 0 0 0 0 4 0 0 0 2 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 [ 28 errs]
  ng 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0103 17 8 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 5 0 1 0 5 1 0 0 [ 44 errs]
   m 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 37144 37 12 0 2 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 1 2 0 [ 97 errs]
   n 0 0 0 0 0 0 1 0 4 0 0 16 0 0 1 0 19 43105 7 1 4 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 6 1 0 0 0 1 1 0 [107 errs]
   N 0 4 7 0 0 0 0 0 2 0 0 0 0 0 0 0 6 1 2323 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 10 1 0 0 [ 35 errs]
   r 0 0 1 1 2 0 0 2 3 0 0 0 8 0 1 1 3 6 4 2386 9 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5 0 0 33 1 0 12 8 2 0 [108 errs]
   w 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 7 0 0 0 0 26 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 [ 10 errs]
   y 0 0 1 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 33 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 3 0 0 0 1 0 0 [ 11 errs]
   s 0 0 1 0 0 0 0 3 0 2 0 0 0 0 0 0 0 3 0 0 0 0 0294 7 1 0 0 3 28 4 27 3 0 5 0 0 0 0 0 0 0 0 0 0 1 1 1 2 [ 92 errs]
  sh 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0189 0 1 0 10 0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [ 14 errs]
   h 0 0 0 0 0 0 0 0 0 18 0 0 0 0 0 0 0 0 0 0 0 0 0 2 13109 0 4 12 3 0 0 0 0 3 0 3 0 0 0 0 0 0 0 1 0 0 0 2 [ 61 errs]
   z 0 1 0 0 0 0 0 0 1 0 1 2 3 0 0 1 0 0 2 0 0 0 0 0 0 0103 0 0 1 0 0 0 15 0 0 0 0 0 0 0 0 1 0 1 0 0 0 3 [ 32 errs]
 ch1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 4 11 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 [ 16 errs]
 ch2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 51 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  1 errs]
 ts1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 2 40 2 1 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 1 1 [ 12 errs]
 ts2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 12 0 1 0 0 0 15137 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [ 32 errs]
  sy 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5 1 0 0 0 0 2 0 66 2 0 5 2 19 0 0 0 0 0 0 0 0 0 4 0 0 [ 41 errs]
  hy 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 8 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  1 errs]
  zy 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 59 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 [  3 errs]
  cy 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 29 2 11 0 0 0 0 0 0 0 0 0 2 0 0 [ 19 errs]
  py 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [  0 errs]
  ky 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 3 0 0 0 0 0 0 0 0 0 0 2 0 36 0 0 0 0 0 0 0 0 0 0 0 0 [  6 errs]
  by 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 [  2 errs]
 gy1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0 0 0 0 0 0 [  2 errs]
 ngy 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 0 0 0 0 0 0 0 0 [  4 errs]
  my 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 [  1 errs]
  ny 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 3 0 0 0 0 0 0 0 [  1 errs]
  ry 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 17 0 0 0 0 0 0 [  6 errs]
  aa 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 7 0 0 0 0 2 [  3 errs]
  ii 0 2 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 43 0 0 0 0 [  3 errs]
  uu 0 1 5 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 77 5 0 4 [ 20 errs]
 eei 0 0 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0118 0 1 [  5 errs]
 oou 0 0 2 0 13 0 0 0 1 0 0 0 0 0 0 0 0 0 0 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0349 3 [ 23 errs]
 sil 0 0 0 0 0 0 28 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 03822 [ 29 errs]
... done
End of Stage 6
```

```
STANDARD SPEAKER
TEST DATA 1702 WORDS
FUZZINESS 1.1
TEMPLATES 25
MIXTURES 12
------------------------ Overall Results -------------------------
PHRASE: %Correct=9.62 [H=198, S=1861, N=2059]
PHONE:  %Corr=81.18, Acc=68.02 [H=12071, D=171, S=2628, I=1957, N=14870]
------------------------------------------------------------------
Confusion Matrix
       a  i  u  e  o  p  p  t  t  k  k  b  b  d  d  g  n  m  n  N  r  w  y  s  s  h  z  c  c  t  t  s  h  z  c  p  k  b  g  n  m  n  r  a  i  u  e  o  s
                      1  2  1  2  1  2  1  2  1  2  1  g                    h        h  h  s  s  y  y  y  y  y  y  y  y  g  y  y  y  a  i  u  e  o  i
                                                                                     1  2  1  2                          1  y                    i  u  l
    a 1280  0  0  2  8  0  0  0  0  0  0  0  0  0  0  0  0  2  0  0  0  3  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 52  0  0  0  0  2  [ 70 errs]
    i    0726  2  1  0  0  0  0  1  0  0  0  0  0  0  0  4  0  0  8  0  2 21  0  1  0  0  0  3  0  1  1  0  0  0  0  7  0  0  0  0  0  3  0132 20 23  0  7  [237 errs]
    u    1 15743 19 14  0  0  0  0  1  1  0  1  0  0  0  3 10  1 68  3  0  3  0  1  0  0  0  0  2  4  1  0  0  0  0  5  0  0  0  0  2  0  0  1189 14  6  4  [369 errs]
    e    0  3  2297  0  0  0  0  1  1  0  0  0  0  0  0  1  0  1  1  0  0  1  0  1  0  1  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9153  0 10  [186 errs]
    o    2  0  2  1448  0  0  0  0  0  1  1  0  0  0  0  1  1  1  0  1  1  0  0  0  1  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0129  3  [146 errs]
   p1    0  0  0  0  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  2 errs]
   p2    0  0  0  0  0  4  0  0 12  1  1  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [ 19 errs]
   t1    0  0  0  0  0  0  0138 28  7  5  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [ 40 errs]
   t2    0  0  0  0  0  5  0  0171  0  2  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  3  [ 11 errs]
   k1    0  0  0  0  0  1  0 21  1318  7  0  0  0  0  0  0  0  0  0  0  0  0  0  1  6  0  0  6  0  2  0  0  0  0  0 20  0  0  0  0  0  0  0  0  1  1  0  1  [ 68 errs]
   k2    1  1  0  0  0  1  0 62 73 74320  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  1  0  2  0  5  0  0  0  2  0 34  0  0  0  0  0  0  1  1  2  1  0  4  [274 errs]
   b1    0  0  0  0  0  0  0  0  1  0  0 10  0  0  2  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  4 errs]
   b2    0  0  0  0  0  0  0  0  1  0  0  1 88  1 25  0  0  1  3  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  [ 35 errs]
   d1    0  0  0  0  0  0  0  0  0  0  0  3  0 50  2  1  0  1  0  0  1  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  9 errs]
   d2    0  0  0  0  0  0  0  0  0  0  0  1 15  1 64  0  0  1  0  0  4  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  [ 23 errs]
   g1    0  0  0  0  0  0  0  3  0  0  0  5  1  6  2 30  0  1  0  0  0  0  1  0  0  0  3  0  0  0  0  0  1  0  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  [ 25 errs]
   ng    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0103 18 10  5  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  5  0  0  1  2  0  0  0  [ 43 errs]
    m    0  0  2  1  1  0  0  0  0  0  0  0  0  0  0  0 42151 30  6  2  1  1  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  1  0  1  0  [ 91 errs]
    n    0  0  0  0  0  0  0  0  3  0  0  1  3  0  1  1 21 37128  4  4  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  3  0  0  0  0  0  1  0  [ 81 errs]
    N    0  0  9  0  0  0  0  0  1  0  0  0  0  0  0  0  4  2  0332  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  9  1  0  0  [ 27 errs]
    r    2  1  2  4  2  0  0  1  0  0  0  0  6  0  1  1  2  2  6  0391  5  6  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 32  0  0  8  4  4  1  [ 90 errs]
    w    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  0  0  0 30  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  5 errs]
    y    0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 41  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  1  0  0  0  [  4 errs]
    s    0  0  2  0  0  0  0  1  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0283  4  3  0  0  2 25 32 29  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  3  [104 errs]
   sh    0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  1169  0  1  0 27  0  0  3  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  [ 34 errs]
    h    0  0  0  0  0  0  0  0  0  6  0  0  0  0  0  0  0  0  0  0  0  0  0  0 10128  0  1 10  1  1  0  0  0  0  0 10  0  0  0  0  0  0  0  1  1  0  0  1  [ 42 errs]
    z    0  0  1  0  0  0  0  0  0  1  1  1  2  0  0  1  0  0  2  0  0  0  1  0  0  0117  0  0  0  0  0  0  5  0  0  0  0  0  0  0  0  0  0  2  0  0  0  1  [ 18 errs]
  ch1    0  0  0  0  0  0  0  0  2  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1 14  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [ 17 errs]
  ch2    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  0  0 43  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  2  0  0  0  1  [  8 errs]
  ts1    0  0  0  0  0  0  0  0  2  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  3 19 18  0  0  0  3  0  2  0  0  0  0  0  0  0  0  0  0  0  1  [ 30 errs]
  ts2    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  3166  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  5 errs]
   sy    0  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  2  2  1  0  0  0  0  2  0 81  0  0  7  1  7  0  0  0  0  0  0  0  0  0  1  0  0  [ 25 errs]
   hy    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  1  5  0  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  [  4 errs]
   zy    0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  0  0  1  0  0  0  0  0  0 56  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  6 errs]
   cy    0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0 31  4  8  0  0  0  0  0  0  0  0  0  2  0  0  [ 17 errs]
   py    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  0 errs]
   ky    0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  2  0  0  0  0  0  0  0  0  3  0 34  0  0  0  0  0  0  0  1  1  0  0  0  [  8 errs]
   by    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  [  2 errs]
  gy1    0  1  1  1  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  [  5 errs]
  ngy    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  2  0  0  0  [  6 errs]
   my    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  [  2 errs]
   ny    0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  3  1  0  0  0  0  0  0  [  1 errs]
   ry    0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  4  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 18  0  0  0  0  0  0  [  5 errs]
   aa    5  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  5  0  0  0  0  0  [  5 errs]
   ii    0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 45  0  0  0  0  [  1 errs]
   uu    0  0  4  2  0  0  0  0  2  1  0  0  0  0  0  0  0  0  3  0  0  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 81  1  0  3  [ 17 errs]
  eei    0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0121  0  0  [  2 errs]
  oou    0  0  1  0 19  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0347  2  [ 25 errs]
  sil    0  0  0  0  0  5  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  1  0  0  0  0  03841  [  6 errs]
... done
End of Stage 6
```

STANDARD SPEAKER
TEST DATA 2693 WORDS
FUZZINESS : 1.6
TEMPLATES : 25

------------------------ Overall Results ------------------------

PHRASE: %Correct=7.84 [H=211, S=2482, N=2693]
PHONE: %Corr=81.53, Acc=67.05 [H=15919, D=205, S=3402, I=2826, N=19526]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	1440	0	0	1	13	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	207	0	0	0	0	6	[237 errs]
i	0	904	0	17	0	3	0	0	0	0	0	0	0	0	0	0	1	1	0	8	0	0	31	0	2	0	2	0	0	0	0	0	1	0	2	0	5	0	0	3	2	4	7	0	206	21	5	0	14	[335 errs]
u	1	17	1080	37	30	2	0	0	0	4	3	0	1	0	0	0	5	12	0	53	6	2	1	0	3	0	3	0	3	5	7	0	0	1	1	2	0	0	0	1	2	1	3	0	4	116	12	16	8	[362 errs]
e	0	15	8	416	0	5	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	3	0	2	0	2	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	1	0	0	6	151	1	2	[201 errs]
o	4	0	2	0	698	3	0	0	0	1	4	0	0	0	0	0	0	5	1	1	1	2	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	119	5	[153 errs]
p1	0	0	0	0	0	1	0	2	4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[11 errs]
p2	0	0	0	0	0	14	0	1	8	1	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[28 errs]
t1	0	0	0	0	0	0	0	153	25	4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[30 errs]
t2	0	0	0	0	0	42	0	2	185	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	[51 errs]
k1	0	0	0	0	0	0	0	30	2	268	26	0	0	0	0	0	0	0	0	0	0	0	0	0	2	15	0	1	4	2	0	0	0	0	1	1	34	0	0	0	0	0	0	0	0	0	3	0	1	[122 errs]
k2	3	0	0	0	0	14	0	56	61	75	444	0	0	0	0	0	0	0	0	0	1	0	4	2	1	3	1	0	3	0	1	0	0	0	4	4	65	0	0	1	0	0	1	3	6	1	3	2	3	[318 errs]
b1	0	0	0	0	0	0	0	0	1	0	0	39	2	6	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	[13 errs]
b2	0	0	0	0	0	0	0	0	0	1	0	0	124	1	20	0	1	0	1	0	2	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	[31 errs]
d1	0	0	0	0	0	0	0	0	0	0	0	0	0	62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	[1 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	13	1	107	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	[17 errs]
g1	0	0	0	0	0	0	0	3	0	0	0	5	0	8	0	36	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	[21 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	146	13	4	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	1	10	0	0	0	0	0	0	0	[43 errs]
m	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	68	246	83	1	10	23	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	2	5	11	2	1	1	0	2	0	2	[217 errs]
n	0	1	0	0	2	2	0	0	0	0	0	2	1	0	0	0	33	34	140	3	14	9	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	1	0	[111 errs]
N	0	5	8	1	0	0	0	0	0	0	1	0	0	0	0	0	7	0	0	423	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	2	0	1	1	4	0	0	0	[38 errs]
r	0	2	1	1	3	0	0	0	2	0	0	4	9	4	4	1	4	9	1	0	563	27	4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	33	1	1	6	9	3	0	[131 errs]
w	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2	0	0	0	75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
y	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	139	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	3	0	0	1	0	0	0	[10 errs]
s	0	0	1	0	0	1	0	3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	380	4	0	0	0	0	41	9	1	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[65 errs]
sh	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	224	0	0	0	10	6	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[17 errs]
h	0	1	1	0	0	1	0	1	0	18	6	0	0	0	0	0	0	0	0	0	0	0	0	1	16	221	0	2	7	0	0	0	3	0	1	0	4	0	0	0	0	0	0	1	0	0	0	1	1	[65 errs]
z	0	1	0	0	0	0	0	0	1	0	2	0	8	1	0	0	0	1	1	0	0	1	0	0	0	1	148	0	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	[23 errs]
ch1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	[13 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	8	64	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	[11 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	39	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	[12 errs]
ts2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	13	188	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[16 errs]
sy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	1	0	0	0	0	0	3	0	80	11	0	16	0	0	0	0	1	0	0	0	0	1	0	0	0	0	[42 errs]
hy	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	59	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	[17 errs]
cy	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	49	0	1	0	0	0	0	0	0	0	3	0	0	0	0	[9 errs]
py	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	[3 errs]
ky	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	4	1	43	0	0	0	0	0	0	0	0	1	0	0	0	[10 errs]
by	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	4	0	0	0	0	0	0	[6 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	[3 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	0	1	0	0	0	0	0	0	[7 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	1	0	0	0	0	0	0	[4 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6	0	0	1	0	0	0	0	[3 errs]
ry	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	37	0	0	0	0	0	0	[9 errs]
aa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	[0 errs]
ii	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	59	0	0	0	0	[1 errs]
uu	0	2	12	2	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	106	4	2	2	[40 errs]
eei	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	155	0	0	[10 errs]
oou	0	0	1	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	474	5	[34 errs]
sil	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5038	[3 errs]

... done
End of Stage 6

Jan 28 19:22 1993 /NFS/atrp10/p10/users/xdonescu/HP/HTK/MHT.train6m.1.6.new Page 1

STANDARD SPEAKER
TEST DATA : 2693 words
Fuzziness : 1.1
Templates : 25
------------------------ Overall Results ------------------------
PHRASE: %Correct=7.69 [H=207, S=2486, N=2693]
PHONE: %Corr=83.47, Acc=67.43 [H=16299, D=192, S=3035, I=3133, N=19526]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	1471	0	0	2	18	0	1	0	7	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	171	0	0	0	0	5	[208 errs]
i	0	960	4	6	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	5	0	0	29	0	0	0	1	0	1	0	0	0	0	0	0	0	7	0	0	2	3	3	4	0	180	17	8	0	6	[279 errs]
u	2	17	1076	27	27	0	0	0	1	2	3	0	1	0	0	0	3	16	0	65	7	1	0	0	3	0	1	0	2	5	4	0	0	0	1	1	2	0	0	0	0	1	2	0	2	136	12	17	9	[370 errs]
e	0	6	7	422	0	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	3	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	7	167	0	3	[202 errs]
o	3	0	4	0	705	0	3	0	1	1	1	0	0	0	0	0	0	6	0	1	0	2	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	108	3	[138 errs]
p1	0	0	0	0	0	5	0	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[7 errs]
p2	0	0	0	0	0	0	12	0	12	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	[16 errs]
t1	0	0	0	0	0	1	0	160	14	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[23 errs]
t2	0	0	0	0	0	0	17	0	210	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[26 errs]
k1	0	0	0	0	0	2	1	21	0	318	9	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	0	1	3	0	1	0	0	0	1	0	26	0	0	0	0	0	0	0	0	0	0	0	0	[73 errs]
k2	1	0	0	0	0	1	12	90	20	59	542	0	0	0	0	0	0	0	0	0	2	0	0	1	1	2	0	0	1	0	0	0	0	0	2	7	16	0	0	0	0	0	1	1	0	0	1	1	0	[219 errs]
b1	0	0	0	0	0	0	1	0	0	0	0	47	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	[6 errs]
b2	0	0	0	0	0	0	0	0	0	1	0	1	126	0	18	0	0	0	1	0	3	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	[31 errs]
d1	0	0	0	0	0	0	0	0	0	0	0	1	0	58	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	[5 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	13	1	105	0	0	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[19 errs]
g1	0	0	0	0	0	0	0	2	0	0	0	4	0	6	0	44	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[13 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	151	10	7	3	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	8	0	0	0	1	0	0	0	[37 errs]
m	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	63	247	105	3	11	13	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6	2	0	0	0	1	2	3	[214 errs]
n	0	0	0	0	1	0	1	0	0	0	0	3	2	0	0	0	18	34	166	2	8	7	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1	0	0	1	0	0	0	[86 errs]
N	0	6	14	0	0	0	0	0	0	0	1	0	0	0	0	0	6	0	0	428	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	5	0	0	1	[36 errs]
r	1	1	1	2	3	0	1	0	0	3	0	7	6	0	1	6	4	7	2	0	579	12	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	14	2	3	1	[108 errs]
w	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
y	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	142	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2	1	0	2	0	0	0	0	0	0	[9 errs]
s	0	0	0	0	0	0	0	3	1	2	0	0	0	0	0	0	1	0	0	0	0	0	0	388	4	0	0	0	2	24	15	2	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	[59 errs]
sh	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	221	0	0	1	14	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[20 errs]
h	0	0	0	0	0	1	0	1	0	55	6	0	0	0	0	0	0	0	0	1	1	0	1	1	12	195	0	0	5	0	2	0	3	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	1	[96 errs]
z	0	0	0	0	0	0	0	0	1	0	3	0	6	1	0	0	0	0	2	0	0	0	0	0	0	0	155	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	[17 errs]
ch1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[19 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	4	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
ts1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	3	30	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[23 errs]
ts2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	10	195	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[11 errs]
sy	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	1	91	2	0	10	0	10	0	0	0	0	0	2	0	0	0	0	0	0	[31 errs]
hy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
zy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	68	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	[8 errs]
cy	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	46	0	5	0	0	0	0	0	0	0	0	0	0	0	0	[12 errs]
py	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	[2 errs]
ky	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3	0	1	2	42	0	0	0	0	0	1	0	1	0	0	1	0	[11 errs]
by	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	4	0	0	0	0	0	0	[6 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	[4 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	0	1	0	0	0	0	0	0	[6 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0	[4 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	0	0	1	0	0	0	0	[3 errs]
ry	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	38	0	0	0	0	1	0	[8 errs]
aa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	[0 errs]
ii	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	[0 errs]
uu	0	0	17	3	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	112	5	2	0	[34 errs]
eei	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	155	0	1	[10 errs]
oou	0	0	1	0	22	0	0	0	0	0	1	0	0	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	478	0	[30 errs]
sil	0	0	0	0	0	0	18	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	5021	[21 errs]

... done
End of Stage 6

Jan 29 10:09 1993 /NFS/atrp10/p10/users/xdonescu/HP/HTK/train8m.1.1_6 Page 1

STANDARD SPEAKER
TEST DATA 2589 words
FUZZINESS : 1.6
TEMPLATES : 25

------------------------ Overall Results ------------------------

PHRASE: %Correct=8.11 [H=210, S=2379, N=2589]
PHONE: %Corr=84.35, Acc=68.60 [H=15771, D=181, S=2746, I=2945, N=18698]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	1466	0	0	1	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	122	0	0	0	0	4	[141 errs]
i	0	921	3	8	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	9	0	0	29	0	0	0	0	0	3	0	0	2	0	0	0	0	5	0	0	0	3	5	5	0	162	27	10	1	7	[281 errs]
u	1	12	1050	25	35	0	0	0	0	0	2	0	1	0	0	0	4	9	1	54	7	0	0	0	1	0	2	0	2	5	10	0	0	0	1	0	0	0	0	0	0	0	0	0	2	150	13	10	8	[355 errs]
e	0	5	4	398	0	0	0	1	2	0	0	0	0	0	0	0	1	0	0	3	0	0	1	0	0	1	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	5	163	0	2	[192 errs]
o	4	0	0	2	685	1	0	0	4	0	2	0	0	0	0	0	0	1	1	0	0	2	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	93	4	[117 errs]
p1	0	0	0	0	0	1	0	2	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[6 errs]
p2	0	0	0	0	0	8	0	0	10	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	[22 errs]
t1	0	0	0	0	0	0	0	155	15	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[22 errs]
t2	0	0	0	1	0	20	0	0	202	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[26 errs]
k1	1	0	1	0	0	0	0	18	2	302	12	0	0	0	0	0	0	0	0	0	0	0	0	0	2	16	0	0	2	0	0	0	0	0	1	0	24	0	0	0	0	0	0	0	1	0	0	0	3	[83 errs]
k2	0	0	0	0	0	5	0	78	34	42	518	0	0	0	0	0	0	0	0	0	0	0	1	1	0	7	0	0	2	0	1	0	0	0	2	0	41	0	0	0	0	0	1	2	2	0	2	1	2	[224 errs]
b1	0	0	0	0	0	1	0	0	0	0	0	46	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
b2	0	0	0	0	0	0	0	0	0	1	0	0	125	1	14	0	1	0	1	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	[23 errs]
d1	0	0	0	0	0	0	0	0	0	0	0	1	0	55	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	13	0	98	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	[19 errs]
g1	0	0	0	0	0	0	0	2	0	0	0	3	0	3	0	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	153	9	5	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	7	0	0	1	2	0	0	0	[33 errs]
m	0	0	0	0	1	0	0	0	2	0	0	1	0	0	0	0	55	276	76	3	5	4	1	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	3	0	9	0	0	1	0	1	1	3	[169 errs]
n	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	15	33	163	1	7	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	1	[78 errs]
N	0	4	17	1	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	415	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	4	1	0	0	[36 errs]
r	1	0	2	1	2	0	0	0	2	1	0	8	8	0	3	4	3	2	2	0	563	6	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	35	0	2	7	2	3	0	[99 errs]
w	0	0	0	0	0	3	0	0	1	0	0	0	0	0	0	0	0	3	0	0	0	66	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
y	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	136	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	3	0	1	0	0	0	0	[10 errs]
s	0	0	0	0	0	0	0	1	1	3	1	0	0	0	0	0	0	1	0	0	1	0	0	384	4	0	0	0	0	23	8	1	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	[50 errs]
sh	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	224	0	0	1	8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[11 errs]
h	0	0	0	0	0	0	0	1	0	38	4	0	0	0	0	0	0	0	0	2	0	0	0	0	6	222	0	0	3	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	1	0	0	0	1	[61 errs]
z	0	0	0	0	0	1	0	0	0	0	1	0	3	1	0	0	0	0	1	0	0	0	0	0	0	1	151	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	[12 errs]
ch1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	14	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[16 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[4 errs]
ts1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	1	2	17	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	[31 errs]
ts2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	7	193	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
sy	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1	0	0	0	0	0	1	91	3	0	5	0	7	0	0	0	0	0	0	0	0	0	0	0	0	[25 errs]
hy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
zy	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	63	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[6 errs]
cy	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	0	2	0	0	0	0	0	1	0	1	0	0	0	0	[8 errs]
py	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[0 errs]
ky	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	1	41	0	0	0	0	0	0	0	1	0	0	1	0	[6 errs]
by	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	[3 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	[3 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	0	0	0	0	0	0	[4 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	[2 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	[1 errs]
ry	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	35	0	0	0	0	0	0	[4 errs]
aa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	[0 errs]
ii	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	0	0	0	0	[1 errs]
uu	0	0	7	0	1	0	0	0	3	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	106	6	0	0	[22 errs]
eei	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	147	0	1	[8 errs]
oou	0	0	1	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	456	1	[23 errs]
sil	0	0	0	0	0	24	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4819	[28 errs]

... done
End of Stage 6

Jan 29 10:30 1993 /p10/users/xdonescu/HTK/standard_speaker/train4m.10.5.1_6 Page 1

STANDARD SPEAKER
TEST DATA : 2693 WORDS
FUZZINESS : 1.6
TEMPLATES : 10
----------------------- Overall Results ------------------------
PHRASE: %Correct=8.95 [H=241, S=2452, N=2693]
PHONE: %Corr=82.33, Acc=68.77 [H=16075, D=251, S=3200, I=2647, N=19526]

Confusion Matrix

	a	i	u	e	o	p1	p2	t1	t2	k1	k2	b1	b2	d1	d2	g1	ng	m	n	N	r	w	y	s	sh	h	z	ch1	ch2	ts1	ts2	sy	hy	zy	cy	py	ky	by	gy1	ngy	my	ny	ry	aa	ii	uu	eei	oou	sil	
a	1481	0	0	1	13	1	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	5	2	1	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	157	0	0	0	0	5	[194 errs]
i	0	943	3	25	0	1	0	0	1	0	0	0	0	0	0	0	1	1	2	8	0	0	36	0	1	0	1	0	3	0	1	0	1	0	0	0	12	0	0	4	2	4	2	0	143	16	7	0	16	[291 errs]
u	4	14	1032	48	46	3	0	1	0	5	4	0	1	0	0	0	3	10	2	31	8	2	1	1	3	3	0	0	2	2	7	0	0	0	1	1	0	0	0	0	1	1	0	0	3	187	5	3	2	[405 errs]
e	0	7	10	445	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1	0	1	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0	0	4	143	1	0	[176 errs]
o	8	0	5	1	712	0	0	1	0	0	3	0	0	0	0	0	0	4	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	104	2	[134 errs]
p1	0	0	0	0	0	0	0	5	3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[12 errs]
p2	0	0	0	0	0	8	0	0	14	2	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[27 errs]
t1	0	0	0	0	0	0	0	159	12	9	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[24 errs]
t2	0	0	0	0	0	17	0	0	210	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[26 errs]
k1	0	0	0	0	0	0	0	31	0	262	36	0	0	0	0	0	0	0	0	0	0	0	0	1	3	12	0	0	3	2	2	0	4	0	0	0	28	0	0	0	0	0	0	0	0	0	1	0	2	[125 errs]
k2	2	0	0	0	1	4	0	95	34	75	519	0	0	0	0	0	0	0	0	0	1	1	0	1	0	4	1	0	1	0	2	0	0	0	1	1	16	0	0	0	0	0	0	0	0	0	0	0	0	[240 errs]
b1	0	0	0	0	0	0	0	0	1	0	0	46	1	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	[7 errs]
b2	0	0	0	0	0	0	0	0	0	1	0	1	120	0	23	0	1	1	2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2	0	0	[37 errs]
d1	0	0	0	0	0	0	0	0	0	0	0	0	0	62	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
d2	0	0	0	0	0	0	0	0	0	0	0	0	15	0	103	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	[21 errs]
g1	0	0	0	1	0	0	0	1	0	0	0	6	0	5	0	42	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	[15 errs]
ng	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	155	5	7	6	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3	5	6	0	0	0	0	0	0	0	[34 errs]
m	0	0	0	1	0	4	0	0	8	0	0	3	1	0	0	0	60	226	83	9	16	29	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	4	8	2	1	0	0	0	0	3	[234 errs]
n	0	0	0	0	1	2	0	0	1	0	0	4	2	0	0	1	22	38	150	2	6	10	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	9	0	0	0	0	0	0	0	[101 errs]
N	1	6	19	0	1	0	0	0	0	0	0	0	0	0	0	0	5	0	0	415	0	0	0	0	4	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	3	0	0	0	[43 errs]
r	0	0	1	2	3	0	0	0	3	0	1	5	2	2	4	3	2	13	10	0	563	14	10	1	0	1	1	0	0	1	0	0	0	0	0	0	4	0	0	0	0	0	17	1	3	2	4	3	0	[113 errs]
w	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	2	0	0	0	70	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
y	0	1	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	139	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	4	0	1	0	0	0	0	[12 errs]
s	0	0	0	0	0	1	0	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	393	2	2	0	0	0	31	6	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	[52 errs]
sh	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	231	0	0	3	4	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[10 errs]
h	0	0	0	0	2	0	0	2	0	30	16	0	0	0	0	0	0	0	0	0	0	0	0	1	5	200	0	0	1	0	4	0	23	0	0	1	3	0	0	0	0	0	0	1	0	0	0	0	0	[89 errs]
z	0	0	0	0	1	0	0	0	2	0	2	0	9	1	3	1	0	0	1	0	3	2	0	0	0	0	143	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	[29 errs]
ch1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[9 errs]
ch2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	9	57	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	[18 errs]
ts1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	34	8	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	[15 errs]
ts2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	13	188	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[16 errs]
sy	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	0	0	0	0	0	0	1	63	1	0	27	0	16	0	0	0	0	0	0	1	0	0	0	0	0	[58 errs]
hy	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[2 errs]
zy	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	1	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	[8 errs]
cy	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	46	0	6	0	0	0	0	0	0	4	0	0	0	0	0	[12 errs]
py	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	[3 errs]
ky	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	2	2	45	0	0	0	0	0	1	0	0	0	0	0	0	[8 errs]
by	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	0	0	0	0	0	0	[5 errs]
gy1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	[3 errs]
ngy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	1	0	0	0	0	0	0	[6 errs]
my	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	0	1	0	0	0	0	0	0	[3 errs]
ny	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1	0	1	0	0	0	0	[3 errs]
ry	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	37	0	0	0	0	0	0	[10 errs]
aa	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	[0 errs]
ii	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	59	0	0	0	0	[1 errs]
uu	0	0	18	4	1	6	0	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	105	2	0	0	[42 errs]
eei	0	3	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	152	0	0	[13 errs]
oou	0	0	1	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	490	1	[17 errs]
sil	0	0	0	0	0	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	5032	[11 errs]

... done
End of Stage 6

Appendix F

Transparent Sheets for Final Talk

SPEAKER-INDEPENDENT FEATURES EXTRACTED FROM A NEURAL NETWORK AND THEIR EVALUATION IN SPEECH RECOGNITION

I. Donescu Y. Kato M. Sugiyama

ATR Interpreting Telephony Research Labs.

2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, JAPAN

Tel: +81-7749-5-1311, Fax: +81-7749-5-1308

Email: xdonescu@atr-la.atr.co.jp

Motivation

- LPC cepstram or FFT spectral power can not be decomposed into speaker characteristics information and the other information
- Success of FPM speaker-independent speech recognition [A Fuzzy Partition Model (FPM) neural network architecture for speaker-independent continuous speech recognition - K.Fukuzawa, Y. Kato and M. Sugiyama 92]
- Expect Speaker-Independent (speaker-robust) features can be extracted from NNs
- NN + HMM Speech Recognition System

  TDNN feature extractor + HMM phoneme recognizer [ATR, Biem91]
- Possible Application to Language Recognition

$\Downarrow$

$\boxed{\text{Speaker-independent Feature Extraction}}$

The proposed algorithm

- Use of a Neural Network to extract features that contain phonemic information, independently of speakers
- Let the number of outputs be a free parameter, independent of the number of phoneme categories

Speaker independent feature space

Arbitrary size of vectors

$\Uparrow$

$\boxed{\text{MAPPING}}$

$\Uparrow$

Speaker-dependent features

Two fundamental problems :

1. What signal target should the network be given ?
2. What criterion should be used to designate the neural network?

The proposed algorithm

→ *A* STANDARD speaker S	→	Target Signal generation
→ *Several* ARBITRARY speakers $\cup_A U^A$	→	Training for speaker-independent feature extraction

- Target signal = Fuzzy Membership Vector
- Criterion : minimize the distance between the sequence of target signals for the standard speaker and the sequence of the actual outputs for any input speaker

The precise algorithm

1. Initialization
   - Calculation of the target signal
   - Initial training of the neural network
2. Alignment between the target signal and the network's output using DTW
   - Calculation of the alignment function
3. Incremental training the neural network f
   - Training for arbitrary speakers
4. Repeat 2–3

Target Signal Generation

- Conventional method for pattern classification :

  phoneme label based :

$$\begin{array}{lcll}
\text{input} & & (/\mathrm{b}/,\ /\mathrm{d}/,\ \ldots /\mathrm{sil}/) & \text{outputs correspond to phoneme labels} \\
\downarrow & & \downarrow & \\
/\mathrm{b}/ & \longrightarrow & (1,\ 0\ \ \ldots\ \ ,0) & \text{target vector}
\end{array}$$

- Proposed method :

  Fuzzy membership function F

$$\begin{array}{lcll}
\text{input} & & \text{output} = \text{Fuzzy membership vector} & \\
\downarrow & & \downarrow & \\
u^{(S)} & \longrightarrow & F(u^{(S)}) = (p_1, ..., p_M) & \text{target vector}
\end{array}$$

$$\Downarrow$$

Templates generated from the set $(U^{(S)})$ of segments for the standard speaker

Calculation of the Fuzzy Membership Vector

2 CASES :

- Category-independent
  - Arbitrary dimension M of the target vector

    $\longrightarrow$ Difficult physical interpretation
- Category-dependent
  - Dimension of the output vector = number of categories = 25

    $\longrightarrow$ Easy physical interpretation

$$\Downarrow$$

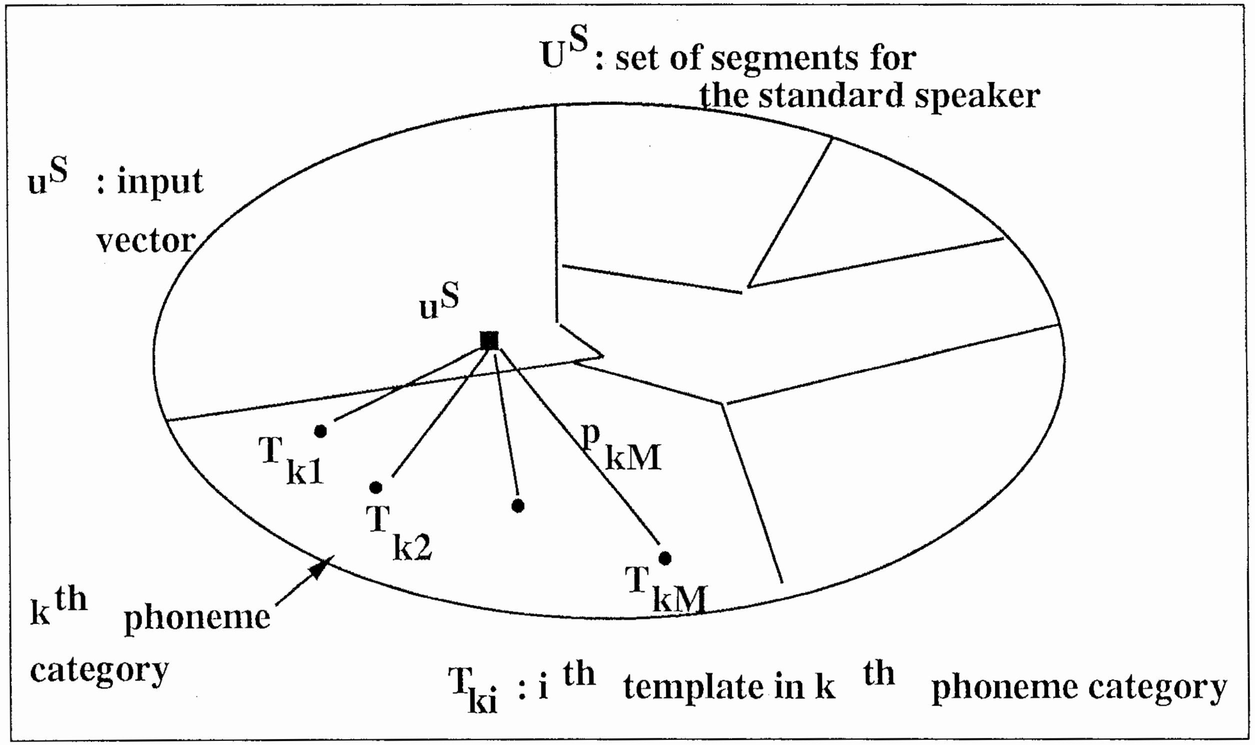
$$\boxed{\text{Category} = \text{phoneme category}}$$

Element of Fuzzy Membership Vector p_k
- Category Dependent Case -

Each phoneme category k is clustered into M templates : $\{T_{k1}, \ldots, T_{kM}\}$

$$p_k = \sum_{l=1}^{M} p_{kl},\ (k = 1, \ldots, K)\ ;\ p_{kl} = 1 \,/\, \sum_{i=1}^{K} \sum_{j=1}^{M} \left(\frac{\|u^S - T_{kl}\|^2}{\|u^S - T_{ij}\|^2} \right)^{1/(g-1)}$$

p_k: kth element of the target vector



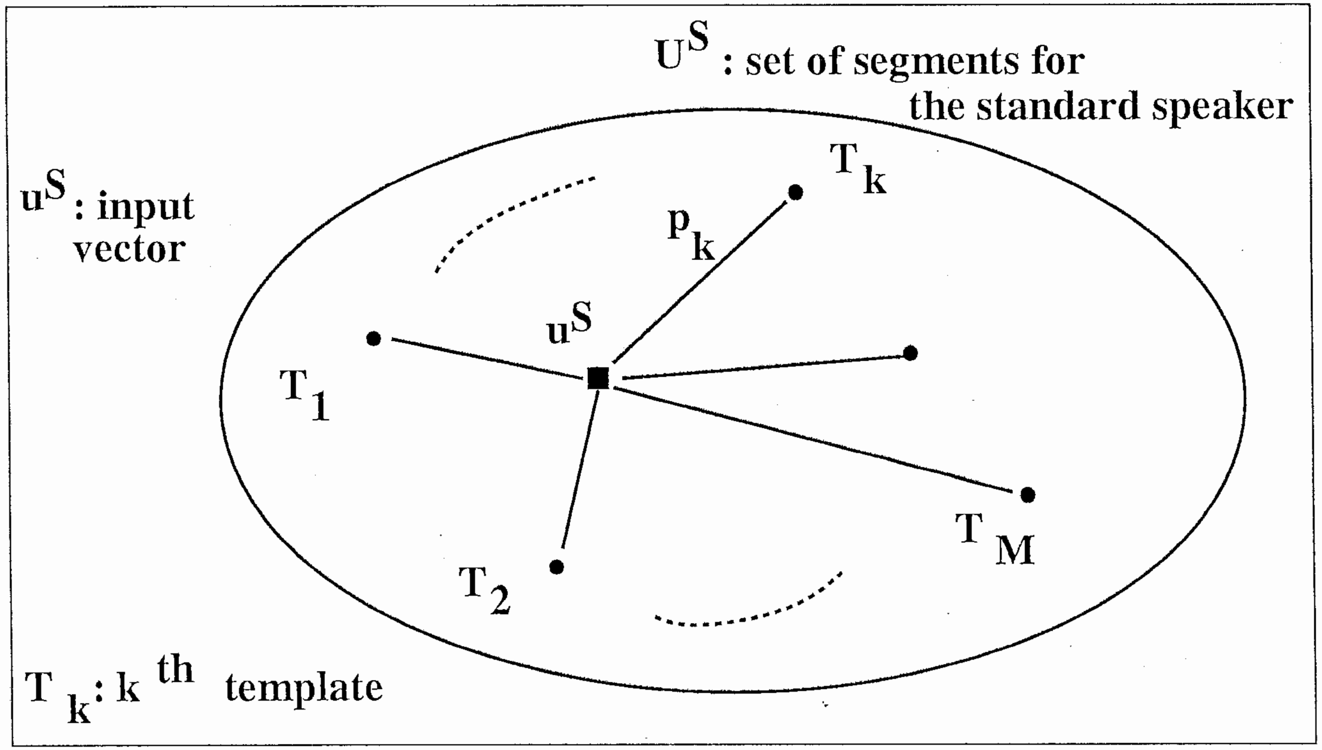
Element of Fuzzy Membership Vector p_k - Category Independent Case -

The whole set $U^{(S)}$ is clustered into M templates : $\{T_1, \ldots, T_M\}$

M arbitrary ;

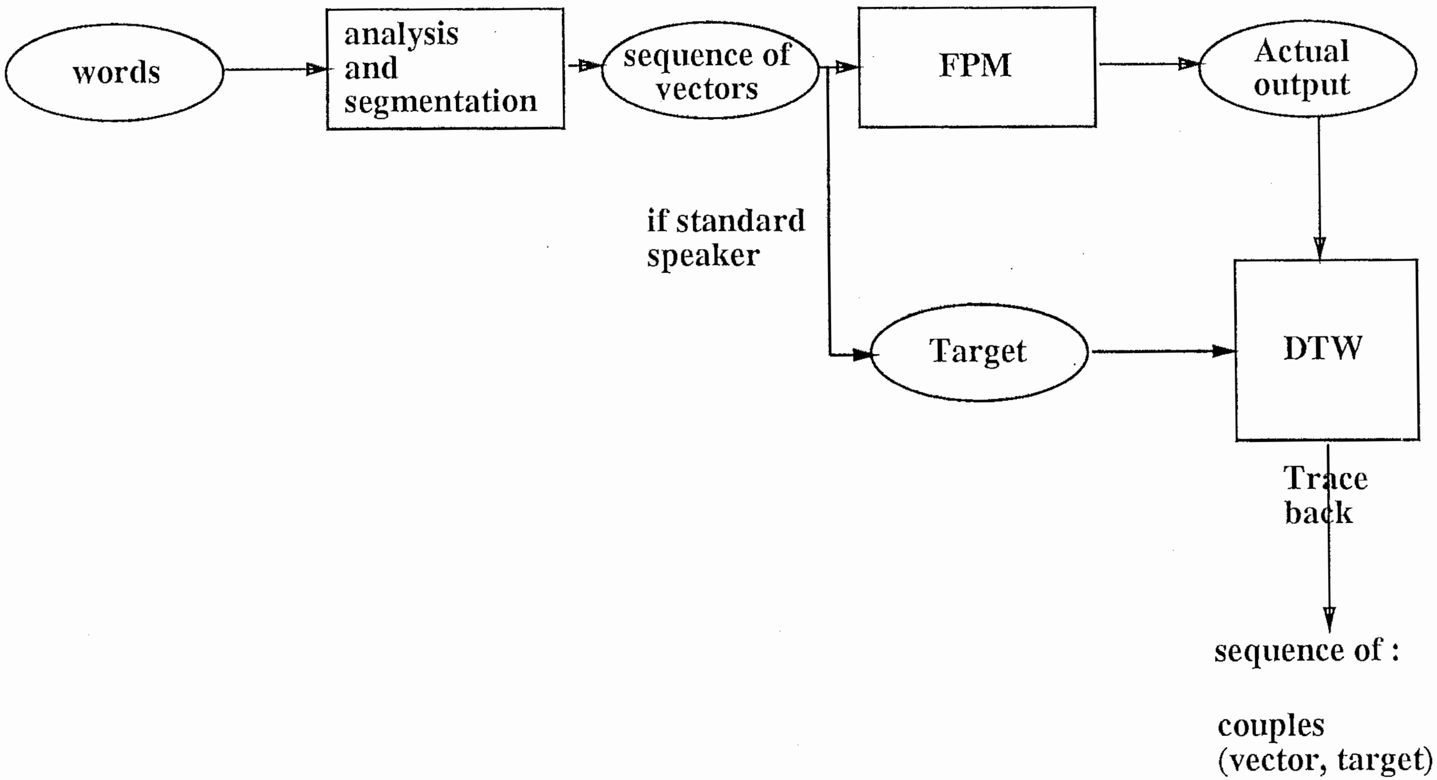
$$p_k = 1 \ / \ \sum_{m=1}^{M} \left(\frac{\|u^S - T_k\|^2}{\|u^S - T_m\|^2} \right)^{1/(g-1)}$$

p_k: kth element of the target vector

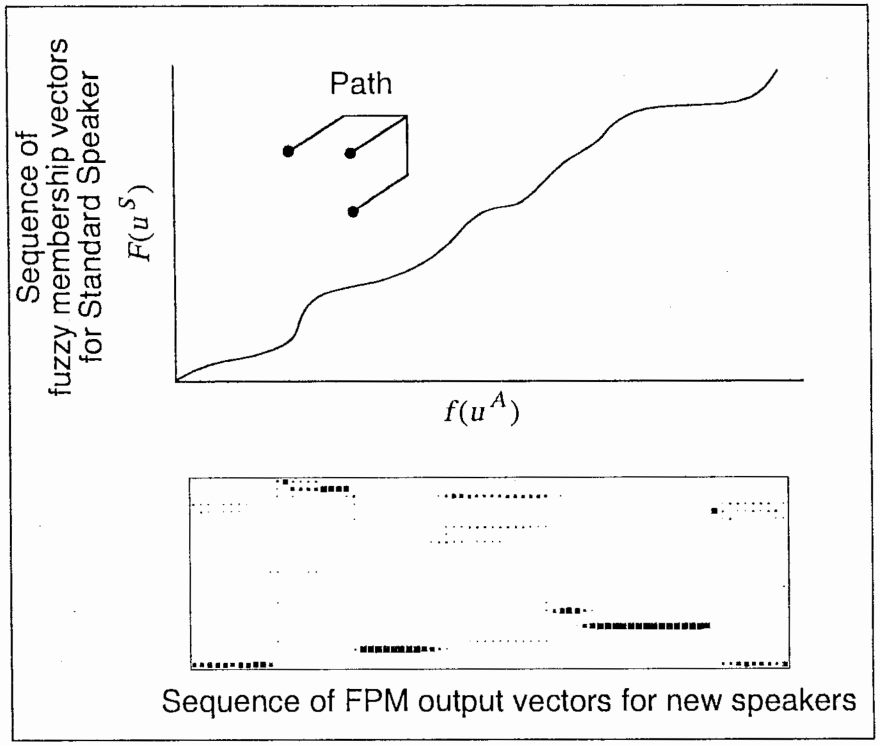


Step 3 : Target generation for incremental training

- Alignement of speech segments (words) = sequence of vectors between the arbitrary speaker A and the standard speaker S



DTW between $F(u^S)$ and $f(u^A)$



Design of Neural Network

New Criterion (Alignment-based)

$$\min_{f} \sum_{u^A_{\tau(i)} \in U^A} d(F(u^S_{\sigma(i)}), f(u^A_{\tau(i)}))$$

$u^A_\tau, u^S_{\sigma(i)}$: Speech segments for arbitrary and standard speakers

σ, τ:Time alignment functions

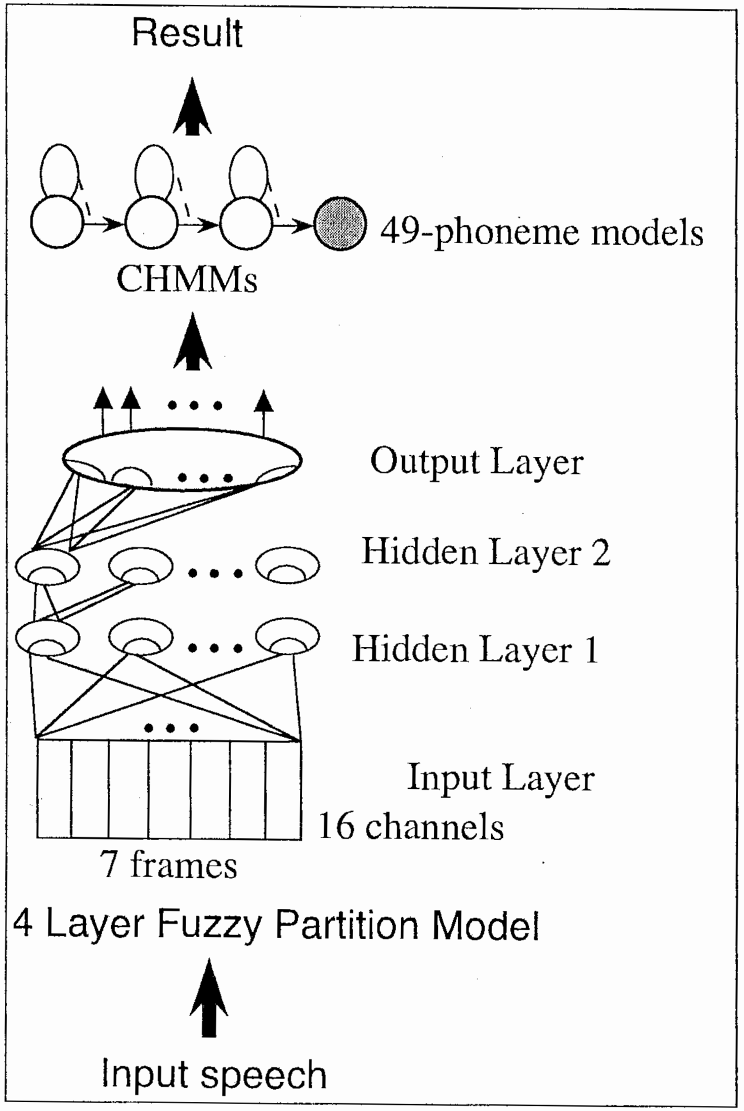
Conventional Criterion (Label-based)

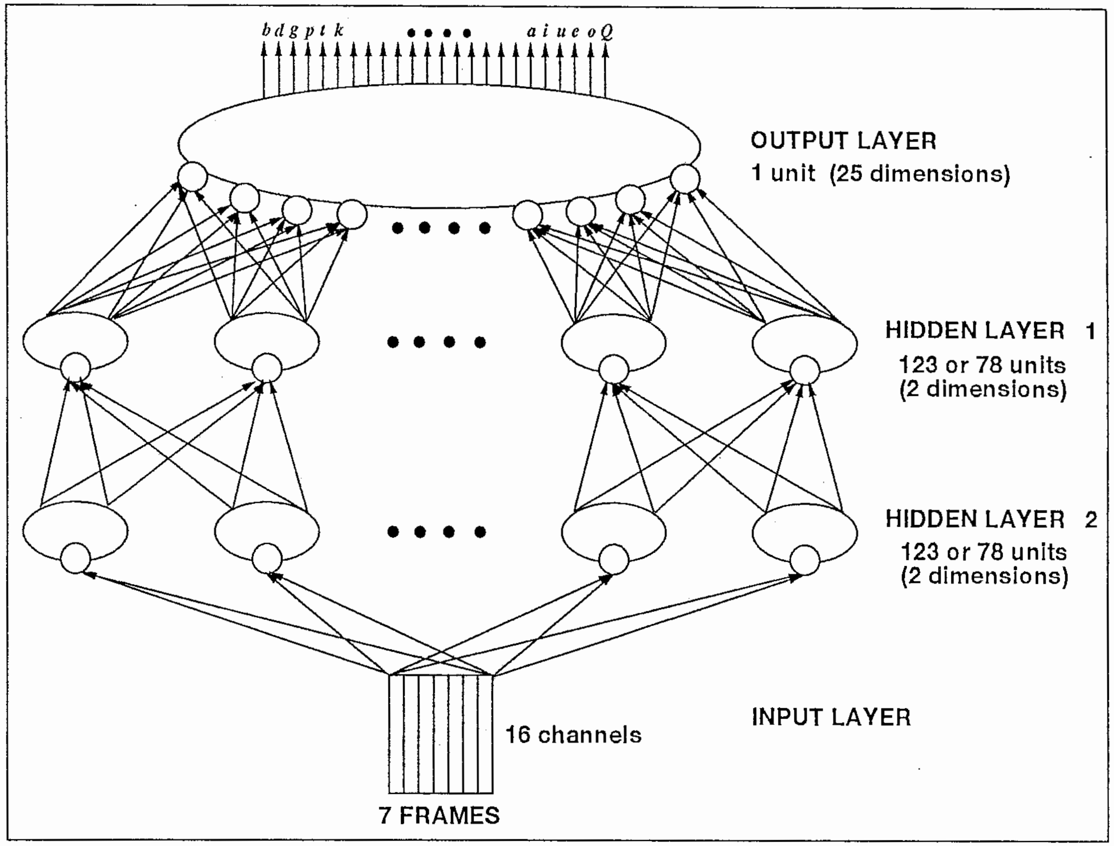
$$\min_{f} \sum_{u^A_i \in U^A} d(F(u^A_i), f(u^A_i))$$

$F(u^A_i) = (\overset{1}{0}\cdots\overset{j}{1}\cdots\overset{M}{0})$: Label

σ, τ: identical

Recognition Process





FPM architecture for Japanese 25 phoneme recognition

Experiments

- Evaluation of the alignment distorsion

$\Downarrow$

Influence of the target generation

- Combination FPM-HMM $\Rightarrow$ recognition performance

$\Downarrow$

Comparison of the methods proposed above :

Conventional pattern classification : Label-based

Category dependent , 8 templates /class

Category dependent, 16 templates /class

Category independent, 25 templates

$\boxed{\textit{Recognition rate} : \text{percentage of labels correctly recognised}}$

- Comparison of FPM-HMM with LPC-HMM
- FPM-LR phrase recognition experiments

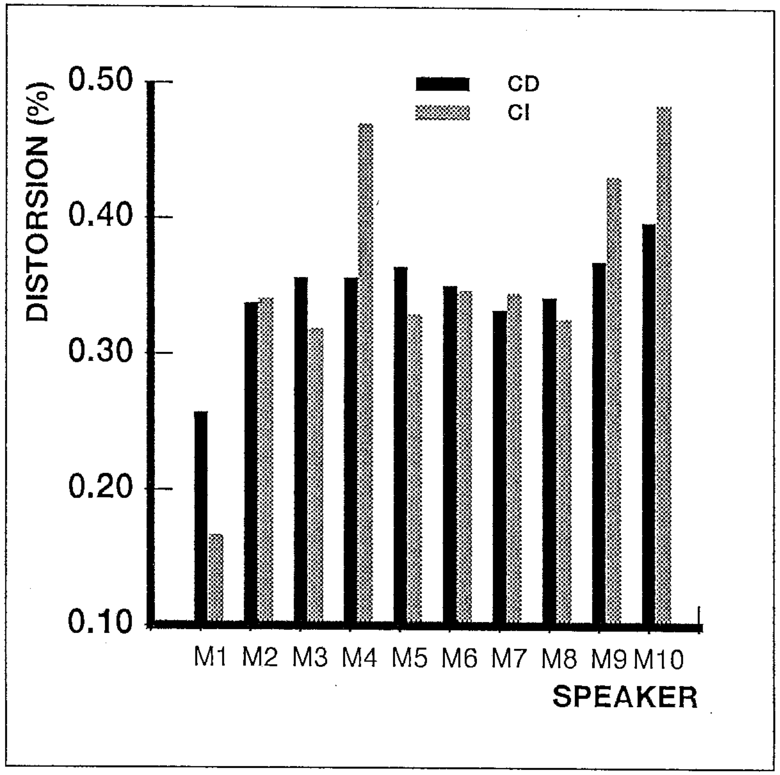
Experimental for speaker-independent FPM training

Neural Network Architecture	
4-layer Fuzzy Partition Model	
Hid. layer 1	123 units, 2 dimensions
Hid. layer 2	123 units, 2 dimensions
Output layer	1 unit, 25 dimensions
Target Signal	
Signal generation	category-dependent generation (8 or 16 templates)/ category-independent generation
Fuzziness	1.1
Training	
Initial training	2,620 isolated Japanese words 50,000 samples
Incremental training	3,624 isolated Japanese words (456 $\times$ 8 speakers) 50,000 samples Includes initial training speaker

Experimental conditions of the recognition part

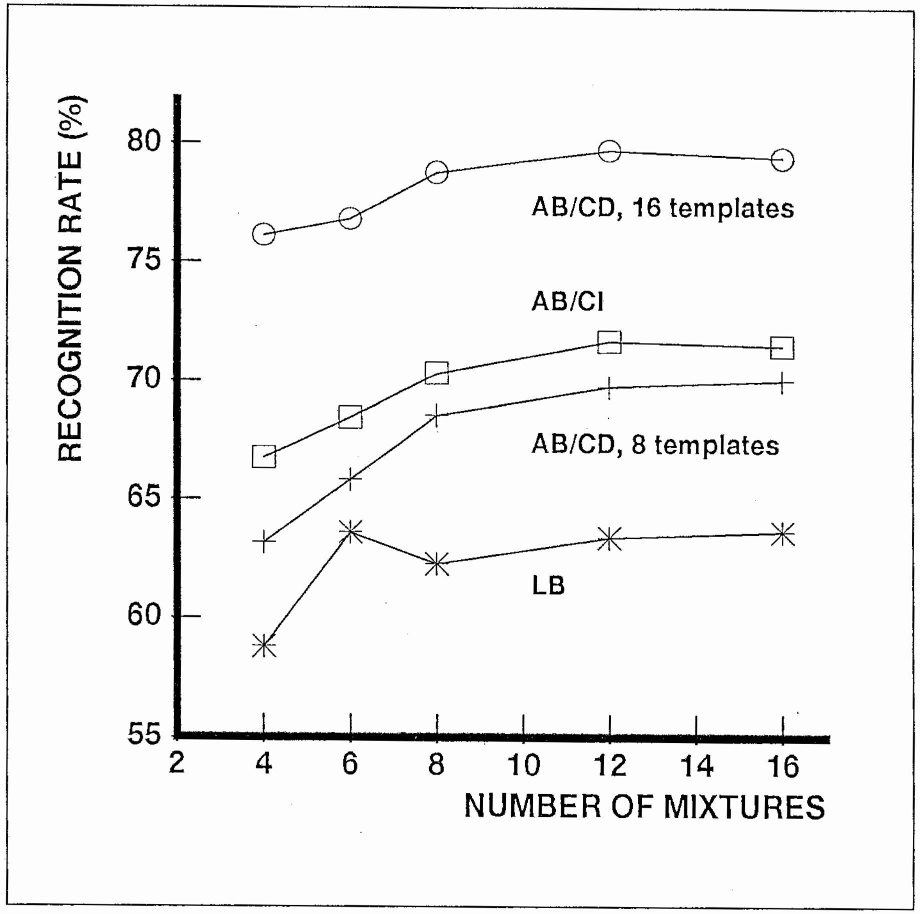
Task	Phoneme recognition
Speakers	Native male Japanese Training: 8 Test: 2
Input parameter	25-dimensional FPM output vector
HMMs	Continuous mixture density type 4-state, 3-loop
Phonemes	48 phonemes + silent class $/p1, p2, t1, t2, k1, k2, b1, b2, d1, d2,\ g,\ ng,\ m,\ n,\ N,\ r,\ w,\ y,$ $s,\ sh,\ h,\ z,\ ch1, ch2, ts1, ts2,\ sy,\ hy,\ zy,\ cy,\ py,\ ky,\ by,\ gy,$ $ngy,\ my,\ ny,\ ry,\ aa,\ ii,\ uu,\ eei,\ oou,\ a,\ i,\ u,\ e,\ o,\ silence/$
Training data	1,728 balanced Japanese words (216 $\times$ 8 speakers)
Test data	432 balanced Japanese words (216 $\times$ 2 speakers)

Alignment distortion for 10 speakers
- Standard Speaker : M1
- Training Speakers : M1-M8
- Open Speakers : M9, M10



Speaker	Distortion	
	category-dep.	category-indep.
Training	0.338	0.330
Open	0.384	0.457

Compared phoneme recognition rates



AB : Alignment-based FPM : { **CI** : Category-independent / **CD** : Category-dependent

LB : Label-based

Analysis conditions for LPC-based feature parameter

Analysis order	14th order
Frequency	12kHz
Frame rate	5ms
Window	256-point (21.3ms) Hamming
Feature parameter	16th LPC cepstrum + 16th Δcepstrum + Power + Δpower

Comparison with LPC-HMMs and FPM-HMMs

Phoneme recognition rate (%)		
FPM-HMMs		LPC-HMMs
AB/CD	AB/CI	LPC
79.4	71.9	83.7

AB/CD: 16 templates/class

AB/CI: 25 templates/class

Experiments on standard speaker

- Value of the fuzziness : 1.1

$\boxed{\text{Fuzzy membership vectors given input for HMMs} \longrightarrow \text{Speaker dependent recognition mode}}$

- Experimental conditions
  - large vocabulary japanese data base (5240 words)
  - 216 balanced japanese words

Training	2692 words
Test	2500 words

STANDARD SPEAKER/CI FPM

Recognition rate in function of the number of mixtures and the number of templates

	Mixtures			
CASE	4	6	8	12
Standard speaker	74.49	76.89	79.33	81.18
Open Speaker, AB/CI	67.85	69.51	70.15	71.91

- New experiments : Fuzziness value : 1.6
- Improvement of the recognition rate for a lower number of mixtures

4 mixtures	81.53
6 mixtures	83.47

Conclusion

- Neural-network based feature extraction
- Evaluation of the robustness of the FPM as a feature extractor by combination with continuous HMM

  -increase of the recognition results when number of templates is larger
- Evaluation of the algorithm by phrase recognition experiments
- The recognition performance is slightly lower with than the LPC-HMM recognition performance, but the data is compressed :

  -25 dimensional training vectors instead of 34 dimensional vectors for LPC

Future studies

- Optimization of the target signal generation : standard speaker
- Improvements possible by changing some of the parameters, for example :
  - increase the fuzziness in order to obtain output values more smoothly distributed
  - increase of the number of templates for the category dependent case
- Optimization of the whole procedure by a global criterion

  e.g Minimun Classification Error or Maximum likelihood