# ADAPTIVE AND DISCRIMINATIVE MODELING FOR IMPROVED MISPRONUNCIATION DETECTION

*Horacio Franco, Luciana Ferrer, and Harry Bratt*

Speech Technology and Research Laboratory, SRI International, Menlo Park, CA

## ABSTRACT

In the context of computer-aided language learning, automatic detection of specific phone mispronunciations by nonnative speakers can be used to provide detailed feedback about specific pronunciation problems. In previous work we found that significant improvements could be achieved, compared to standard approaches that compute posteriors with respect to native models, by explicitly modeling both mispronunciations and correct pronunciations by nonnative speakers. In this work, we extend our approach with the use of model adaptation and discriminative modeling techniques, inspired on methods that have been effective in the area of speaker identification. Two systems were developed, one based on Bayesian adaptation of Gaussian Mixture Models (GMMs), and likelihood-ratio-based detection, and another one based on Support Vector Machines classification of supervectors derived from adapted GMMs. Both systems, and their combination, were evaluated in a phonetically transcribed Spanish database of 130,000 phones uttered in continuous speech sentences by 206 nonnative speakers, showing significant improvements from our previous best system.

***Index Terms***— Mispronunciation detection, computer-aided language learning.

## 1. INTRODUCTION

Using computers to help students learn and practice a new language has long been seen as a promising area for the use of automatic speech recognition (ASR) technology. It could allow spoken language to be used in many ways in language-learning activities, for example by supporting different types of oral practice and enabling feedback on various dimensions of language proficiency, including language use and pronunciation quality. A desirable feature of the use of speech technology for computer-aided language learning (CALL) is the ability to provide meaningful feedback on pronunciation quality. In this area of pronunciation scoring, the smaller the unit to be scored, the higher the uncertainty in the associated score [1]. Currently, the most reliable estimates of pronunciation quality are overall levels obtained from a paragraph composed of several sentences that can be used to characterize the speaker's overall pronunciation proficiency. At this level, it has been shown that automatic scoring performs as well as human scoring [2].

For many CALL applications we would like to score smaller units, to allow the student to focus on specific aspects of his or her speech production. For instance, overall pronunciation scoring can be obtained at the sentence level [3], [4], with a level of accuracy that, while lower than that of human scoring, can nonetheless provide valuable feedback for language learning [5]. More detailed feedback, at the level of individual phones, can direct attention to specific phones that are mispronounced [1], [6-11].

In earlier work [9], we compared two approaches for phone mispronunciation detection. The first was based on using native models as a reference, obtaining a measure of the phone-level degree of match to a corresponding native model [1]. The second approach was based on using explicit acoustic models for the correct and for the mispronounced utterances of a phone, and computing a likelihood ratio using these two models. We found that the second approach was more accurate, resulting in an average 9% relative reduction in the equal error rate (EER) of the mispronunciation detection.

In this paper we extend the work on acoustic modeling of correct and mispronounced nonnative phones by using more advanced acoustic modeling techniques based on adaptation and discriminative modeling. The proposed techniques are inspired by approaches that have been effective on significantly improving accuracy in another area of speech technology – namely, speaker recognition. The first proposed approach uses model adaptation in a form that is inspired by the Gaussian Mixture Model-Universal Background Model (GMM-UBM) speaker verification system proposed by Reynolds et al. [12]. This approach has shown to be more effective than other adaptation approaches when used for a detection task, particularly when limited adaptation data is available. The second proposed approach is based on the use of discriminant classifiers based on support vector machines (SVMs) using as an input feature a GMM supervector consisting of the stacked means and weights of the mixture components. The GMM supervector is obtained by adapting a GMM-UBM to a test utterance [13]. We also explored the combination of the scores from these two approaches.

The outline of this paper is as follows: in Section 2 we review the baseline approach and present the new approaches explored in this paper, in Section 3 we describe the database that we use to evaluate them, and in Section 4 we present our experimental results comparing the different approaches. Section 5 concludes this work.

## 2. PHONE-LEVEL MISPRONUNCIATION DETECTION APPROACHES

Earlier modeling approaches [1], [7] used basically a native model to produce a measure of goodness for the nonnative speech. While this measure correlates very well with human judgments for longer segments (i.e., paragraphs or sentences), the correlation decreases for shorter segments, such as phones [1].

After labeling a database for phone-level mispronunciations [14], in [9] we explored a more detailed acoustic modeling to attempt to capture the subtle differences between the nonnative speech realizations that are considered acceptable versus the nonnative speech realizations that are considered mispronounced. In that work we compared two mispronunciation detection schemes. The first approach was based on phone log-posterior scores [15], [1]. The phone log-posterior scores are similar to the GOP scores introduced in [7], but log-posterior probabilities are computed at the frame level, and averaged over the frames of a given phone segment. The second approach was based on explicit acoustic modeling, using GMMs, of the nonnative productions of correct and mispronounced phones [9]. A log-likelihood ratio (LLR) of mispronounced and correct phone models was used as the measure of pronunciation quality in the second method. We found significant improvements from the use of the explicit mispronunciation modeling using the GMM-LLR approach. Nevertheless, it should be noted that the explicit acoustic modeling of mispronunciation comes at a price: it is necessary to collect and annotate a nonnative training database, and the resulting models are dependent on the first language of the nonnative speakers.

In this work we aim to further develop the explicit modeling of mispronunciations by using newer acoustic modeling techniques that can be more effective dealing with the challenges of this task. In the following approaches we assumed that the phonetic segmentation is obtained by computing a forced alignment with an HMM, using the corresponding speech transcription and a pronunciation dictionary.

### 2.1. System 1: LLR of independently trained GMMs

Our baseline approach for mispronunciation detection is the GMM-LLR proposed in [9], where for each phone class we trained two different GMMs: one model is trained with the "correct" native-like pronunciations of a phone, while the other model is trained with the "mispronounced" or nonnative pronunciations of the same phone. In the evaluation phase, for each phone segment $q_i$, a length-normalized log-likelihood ratio score $LLR(q_i)$ was computed by using the "mispronounced", $\lambda_M$, and the "correct", $\lambda_C$, pronunciation models, respectively, where $LLR(q_i)$ is defined as

$$LLR(q_i) = \frac{1}{d_i} \sum_{t=t_i}^{t_i+d_i-1} [\log P(y_t \mid q_i, \lambda_M) - \log P(y_t \mid q_i, \lambda_C)] \quad (1)$$

where $P(y_t \mid q_i, \lambda)$ is the probability of the acoustic feature $y_t$ for the frame at time $t$ given the phone class $q_i$ and the model $\lambda$. The normalization by the phone duration $d_i$ allows definition of unique thresholds for the LLR for each phone class, independent of the lengths of the segments. A mispronunciation is detected when the LLR is above a predetermined threshold, specific to each phone. In this system we used diagonal covariance GMMs for all phone models, and for each phone and each class ("correct" or "mispronounced"), the corresponding GMM was created with a number of mixture components proportional to the number of samples for the phone for that class. The proportion is a tunable parameter that can be optimized.

### 2.2. System 2: LLR of adapted GMMs

This system is similar to baseline System 1, except that the models for each class ("correct" and "mispronounced"), for each phone, are obtained by adaptation. The model to which they are adapted is trained using all the samples from a given phone, ignoring the class. We use Bayesian adaptation [16] to adapt this class-independent GMM to all the "correctly" pronounced training examples for a phone, and obtain the adapted "correct" model for such phone. We proceed similarly to obtain the adapted "mispronounced" model for such phone. For these class-dependent GMMs we adapt both the means and the mixture weights to the class-specific data.

With these two adapted models we can compute, for any test phone, the LLR of the adapted "mispronounced" model to the adapted "correct" model. The key point in this approach is that the models used to compute the LLR are not trained independently, but are derived from the same class-independent model. This provides a tighter coupling between the "mispronounced" and the "correct" model, which has been shown to produce better performance in the area of speaker detection [12].

### 2.3 System 3: SVM classifier based on adapted GMM supervector

Inspired by the work on [13], we use the class-independent GMM trained for System 2 to create supervectors by adapting this GMM to each phone instance. The supervector for a certain phone instance is obtained by adapting the means and mixture weights of the original GMM to the acoustic feature vector representing the phone. This operation corresponds to a transformation of the phone segment acoustic feature vectors into a fixed high-dimension feature vector in the GMM supervector space. In our preliminary experimentation, we found that adapting only the means or only the weights of the GMM gives worse performance than adapting both means and weights. The supervectors are then normalized to have the same variance in all dimensions, using the statistics found in the training data, and fed to a linear SVM [17].

Given a test phone instance, its supervector x is first computed by adaptation of the class-independent GMM to the phone feature vector frames. The distance of that supervector to the SVM hyperplane is taken as the score for the phone.

### 2.4. System 4: Combination of Systems 2 and 3

We combined the two best systems by using a simple weighted combination of the scores given by each of the two systems for each phone. No tuning was done on the weights because we lacked an additional development set with which to tune it. We used a weight of 0.25 for System 2 and 0.75 for System 3, since the range of the scores of System 3 is around three times smaller than the range for those of System 2. So, the weight was just used for equalizing the score ranges. The same weight was used for all phone classes for each system.

## 3. DATABASE DESCRIPTION

To evaluate these modeling approaches we used a phonetically transcribed subset of the nonnative Spanish database described in [14]. All speech data was read speech from Spanish newspapers with no repeated sentences, aiming at developing text-independent systems. Four native Spanish-speaking expert phoneticians transcribed 2550 sentences, totaling 130,000 phones, of nonnative speech data. Those sentences, randomly divided among the transcribers, were produced by 206 nonnative speakers whose native language was American English. Their levels of proficiency

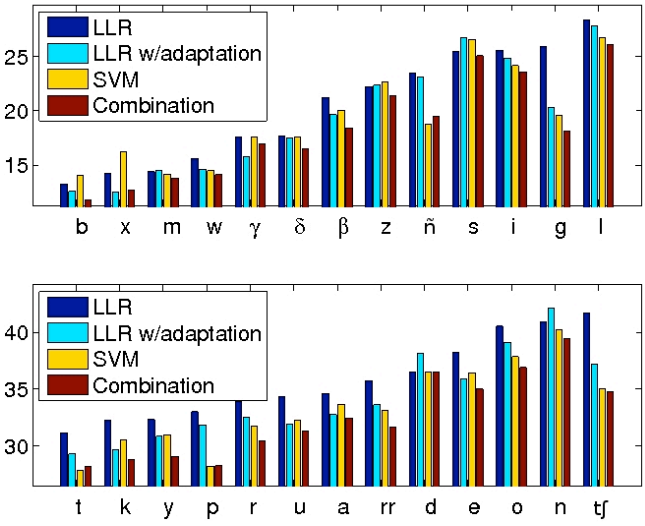


Figure 1: EER in % for the four systems, for each phone, in increasing order of EER for the baseline System 1 (LLR). System 2 is denoted as LLR w/adaptation, System 3 is SVM based, and System 4 is the combination of Sytems 2 and 3.

were varied, and an attempt was made to balance the number of speakers by level of proficiency as well as by gender. An additional set of sentences (one newspaper sentence from each of the 206 speakers), the common pool, was transcribed by all four phoneticians to assess human-human consistency. For this study, the detailed phone-level transcriptions were collapsed into two categories: native-like and nonnative pronunciations.

To assess how consistently humans can detect mispronunciations we used the 206 common sentences from the transcribed database, and used the Kappa coefficient statistic [18], to determine how reliably the transcribers agree on the transcription for each of the 28 native phones. For nine of the phones (/β/, /δ/, /γ/, /b/, /w/, /m/, /ñ/, /i/ ,/s/), all four transcribers showed at least a moderate level of agreement (using $K > 0.4$ to mean "moderate" agreement).

## 4. EXPERIMENTAL SETUP AND RESULTS

Our mispronunciation detection approaches assume that the phonetic segmentation is given and accurate. Therefore, the task for which the mispronunciation detection is used must be designed to ensure a good speech recognition performance. Examples of such tasks are reading aloud and multiple-choice exercises.

For our experiments we generated phonetic alignments using the EduSpeak [19] HMM-based speech recognizer. The acoustic features were standard 13-dimensional Mel frequency cepstral coefficients (MFCCs) plus their delta coefficients obtained every 10 ms, based on a sliding 25-ms Hamming window. The C0 coefficient was normalized by the maximum over each sentence. Cepstral mean normalization was applied at the sentence level for C1 to C12. The acoustic models used to generate the phonetic alignments were gender independent, Genonic GMMs, as introduced in [20].

Given the alignments, the detection of mispronunciation is reduced to a binary classification of the phone's feature vectors, as the phone class is given by the alignments. The performance of the mispronunciation detection algorithms was evaluated as a function of the threshold applied to the detection score, for each phone class. For each threshold we obtained the machine-produced labels "correct" (C) or "mispronounced" (M), for each phone utterance. Then, we compared the machine labels with the labels obtained from the phoneticians' transcriptions. For each threshold we computed two error measures: the probability of a false positive, estimated as the percent of cases where a phone utterance is labeled by the machine as incorrect when it was in fact correct, and the probability of a false negative – that is, the probability that the machine labeled a phone utterance as correct when it was in fact incorrect.

For each phone we evaluated the receiver operating characteristic (ROC) curve, and found the points of equal error rate (EER), where the probability of false positive is equal to the probability of false negative. Note that in an actual application of the mispronunciation detection system, criteria other than the EER may define an operating point along the ROC curve. For instance, for pedagogical reasons we may want to impose a maximum acceptable level of false positives.

We used a four-way jackknifing procedure to train and test these approaches on the same phonetically transcribed nonnative database. We trained models using data from three partitions and tested on the remaining partition, rotating the procedure four times over the four partitions. There were no common speakers across any of the partitions. When reporting results for a given system, the errors obtained for each of the four partitions were pooled to obtain mispronunciation detection average performance on the complete database.

### 4.1 System training and tuning

For System 1 we replicated the results from the original work [9] with our current GMM training software. As different phones have different amounts of training data, we explored the number of mixture components to use for each phone class GMM, and found that the proportion of 25 training samples for each mixture component resulted in the best performance for this system. The number of mixture components ranged from 2 to 531.

In developing System 2 we found that the optimal number of mixture components of the class-independent GMM for each phone was approximately the sum of the sizes of the class-dependent GMMs for each phone from System 1. We estimated the Bayesian adapted model using the detailed procedure described in [12].

To train the SVM for System 3 we used the software package SVM-light [17], which can efficiently handle large training sets and allows for asymmetric cost factors.

### 4.2 Experimental results

In Table 1 and Figure 1 we show the EER for mispronunciation detection for each of the systems and for every phone. In Table1 we also show, for each system, weighted averages of the EER over all phones, where the weight was the relative frequency of each phone.

The GMM approach of System 1, the baseline system, had an average EER of 31.8%. The phones with the best performance were the approximants /β/, /δ/, /γ/, the voiced stop /b/, the semivowel /w/, the nasal /m/, and the fricative /x/. These phones have very good agreement with the phones with the highest Kappa.

The use of adaptation in System 2 produced a significant reduction of EER across most of the phones. The largest improvements occurred for the voiced stop /g/, the voiceless velar

fricative /x/, the voiceless palatal affricate /ʧ/ and /g/, (21.6%, 11.6%, 10.7%, and 10.1% relative reduction with respect to System 1, respectively). These phone classes had less data than the average, showing that the adaptation approach is effective in dealing with smaller amounts of training data. Some of the largest EER reductions occurred for phones that were not necessarily among the most reliably transcribed, which suggests that the adaptation approach took advantage of additional useful information in the noisy transcriptions. The overall weighted EER reduction with the adaptation approach was 3.5% relative to the baseline system. There were also a few phones where the EER was slightly worse than the baseline, mainly for /s/, /d/, and /n/.

Table 1. Equal Error Rate at the phone level for the four Systems studied. Weighted averages of the EER for each system are shown at the bottom. We also show the number of samples labeled as Correct or Mispronounced for each phone.

| | Equal Error Rate | | | | # samples | |
|---|---|---|---|---|---|---|
| Phone | Sys. 1 | Sys. 2 | Sys. 3 | Sys. 4 | Corr. | Mispr. |
| b | 13.3 | 12.6 | 14.1 | 11.8 | 668 | 490 |
| x | 14.2 | 12.6 | 16.3 | 12.7 | 752 | 190 |
| m | 14.4 | 14.5 | 14.2 | 13.9 | 4037 | 874 |
| w | 15.6 | 14.7 | 14.5 | 14.1 | 922 | 628 |
| δ | 17.7 | 17.5 | 17.6 | 16.5 | 1188 | 2497 |
| ɣ | 17.6 | 15.8 | 17.6 | 17.0 | 283 | 788 |
| g | 25.9 | 20.3 | 19.6 | 18.1 | 1124 | 143 |
| β | 21.3 | 19.7 | 20.1 | 18.4 | 554 | 1471 |
| ñ | 23.4 | 23.1 | 18.8 | 19.5 | 1157 | 559 |
| z | 22.3 | 22.4 | 22.7 | 21.4 | 238 | 1246 |
| i | 25.6 | 24.8 | 24.2 | 23.6 | 6153 | 1539 |
| s | 25.4 | 26.7 | 26.6 | 25.1 | 9520 | 587 |
| l | 28.3 | 27.8 | 26.7 | 26.1 | 4358 | 1729 |
| t | 31.1 | 29.3 | 27.9 | 28.1 | 3671 | 1902 |
| p | 33.0 | 31.8 | 28.2 | 28.3 | 2098 | 1310 |
| k | 32.3 | 29.6 | 30.5 | 28.8 | 2142 | 1861 |
| y | 32.3 | 30.8 | 30.9 | 29.1 | 3050 | 730 |
| r | 33.9 | 32.5 | 31.7 | 30.5 | 4561 | 3258 |
| u | 34.3 | 31.9 | 32.3 | 31.3 | 2420 | 592 |
| rr | 35.7 | 33.6 | 33.1 | 31.7 | 613 | 2192 |
| a | 34.6 | 32.7 | 33.6 | 32.4 | 12655 | 2617 |
| ʧ | 41.7 | 37.2 | 35.0 | 34.8 | 510 | 129 |
| e | 38.2 | 35.9 | 36.4 | 35.0 | 13258 | 4385 |
| d | 36.5 | 38.1 | 36.5 | 36.5 | 965 | 107 |
| o | 40.5 | 39.1 | 37.8 | 36.9 | 10072 | 2627 |
| n | 40.9 | 42.1 | 40.3 | 39.4 | 8944 | 601 |
| Avg. | 31.8 | 30.7 | 30.3 | 29.3 | - | - |

The SVM-based System 3 produced a larger average EER reduction than System 2. Compared with System 1, the biggest EER reductions happened in phones /g/, / ñ /, /ʧ/, and /p/ (24%, 20%, 16%, and 14% relative EER reductions, respectively). Comparing Systems 2 and 3, we find that significant additional EER reductions occurred for the phones /ñ/ and /p/ (18.8% and 11.5% relative reductions, respectively), while some other phones had smaller gains. There was also significantly degraded performance for phones /x/, /b/, /ɣ/ (-29.4%, -11.7%, -11.2%) relative to System 2. These performance losses were unexpected, as these phones are among the set of best-performing phones in the baseline system, and they are also among the phones with high consistency in their transcriptions. They do have less training data compared to the other phones in the set of best-performing phones in the baseline system, which might have had an effect on this approach. The overall weighted EER reduction of System 3 with respect to System 1 and System 2 was 4.7% and 1.3%, respectively.

System 4, based on combining the scores of System 2 and System 3, was the best-performing system. It produced improvements over System 3 mostly on the phones /x/ and /b/ that had been degraded by System 3, also having moderate gains over most phones. Comparing System 4 with the baseline System 1 we appreciate that the combination of systems produces large gains on most phones ranging from relative EER reductions of 29.6% for /g/ to ~18% for /ʧ/, /n/, and /p/, and 10% to 12% for /b/, /rr/, /x/, /b/, /t/, /u/, /w/, gradually going down for the remaining phones. It is noteworthy that the EER reductions obtained by System 2 and System 3 with respect to the baseline resulted in reinforcing EER reductions for many phones in the combined system (see, for instance, /g/, /rr/, /ʧ/, etc.) Overall, the weighted EER relative reduction of System 4 with respect to the baseline System 1 was 8%, showing an almost additive combination of the average gains from System 2 and System 3. These results suggest that the improvements brought by these two systems over System 1 are highly complementary. Also, we observe that in a few cases the EER of the combination system was slightly higher than either of the combined systems, suggesting that per-phone combination weights could improve System 4.

In general, the phones with the lowest EER were also those more consistently labeled by the transcribers. Nine phones had EER below 20%; among those, seven had Kappa above 0.4 while the other two did not have enough data in the common sentences data set to assess Kappa.

## 5. CONCLUSION

We studied approaches for detection of mispronunciations, based on the explicit acoustic modeling of correct and mispronounced training examples.

We proposed and analyzed two new mispronunciation detection algorithms. The first is based on computing the GMM likelihood ratio of adapted models to the correct and mispronounced training examples for each phone class. The second approach is based on discriminative modeling using supervectors derived from the parameters of GMMs adapted to phone segments modeled by SVM classifiers trained on examples of correct and mispronounced phones. Both methods proved to be superior to our previous mispronunciation detection system based on the LLR of independently trained GMMs, which had been previously shown to outperform the standard method that uses phone log-posteriors as scores.

The first approach produced a relative reduction of the weighted average EER of 3.5%, while the second approach produced a 4.7% relative reduction of the weighted average EER. Furthermore, a third system based on the combination of the scores provided by the first two new systems produced an almost additive error reduction, resulting in an 8% relative reduction for the average EER.

## 6. REFERENCES


[1] Y. Kim, H. Franco, and L. Neumeyer, "Automatic pronunciation scoring of specific phone segments for language instruction," Proc. EUROSPEECH 97, pp. 649–652, Rhodes, 1997.

[2] J. Bernstein, M. Cohen, H. Murveit, D. Rtischev, and M. Weintraub, "Automatic evaluation and training in English pronunciation," Proc. ICSLP 90, pp. 1185–1188. Kobe, Japan, 1990.

[3] L. Neumeyer, H. Franco, M. Weintraub, and P. Price, "Automatic text-independent pronunciation scoring of foreign language student speech," Proc. ICSLP 96, Philadelphia, Pennsylvania, pp. 1457–1460, 1996.

[4] H. Franco, L. Neumeyer, Y. Kim, and O. Ronen, "Automatic pronunciation scoring for language instruction," Proc. ICASSP 97, pp. 1471–1474, Munich, 1997.

[5] K. Precoda, C. Halverson, and H. Franco, "Effect of speech recognition-based pronunciation feedback on second language pronunciation ability," Proc. InSTILL2000: Integrating Speech Technology in Learning, pp. 102-105. University of Albertay, Dundee, Scotland. 2000.

[6] M. Eskenazi, "Detection of foreign speakers' pronunciation errors for second language training – preliminary results," Proc. ICSLP 96, pp. 1465–1468. 1996.

[7] S. Witt and S. Young, "Language learning based on non-native speech recognition," Proc. EUROSPEECH 97, pp. 633–636, Rhodes, 1997.

[8] O. Ronen, L. Neumeyer, and H. Franco, "Automatic detection of mispronunciation for language instruction," Proc. EUROSPEECH97, pp. 645-648, Rhodes, 1997.

[9] H. Franco, L. Neumeyer, M. Ramos, and H. Bratt, "Automatic detection of phone-level mispronunciations for language learning," Proc. Eurospeech 99, 2, pp. 851-854, Budapest, Hungary, 1999.

[10] K. Truong, A. Neri, C. Cucchiarini, and H. Strik, "Automatic pronunciation error detection: An acoustic-phonetic approach," Proc. InSTIL/ICALL Symposium, 17-19 June, Venice, Italy, pp. 135-138. 2004.

[11] H. Strik, K. Truong, F. de Wet, and C. Cucchiarini, "Comparing classifiers for pronunciation error detection," Proc. Interspeech-2007, Antwerp, Belgium, pp. 1837-1840. 2007.

[12] D.A. Reynolds, T. F. Quatieri, and R.B. Dunn, "Speaker verification using adapted Gaussian mixture models," Digital Signal Processing, 10, pp. 19-41. 2000.

[13] W. Campbell, D. Sturim, D. Reynolds, and A. Solomonoff, "SVM based speaker verification using a GMM supervector kernel and NAP variability compensation," ICASSP 2006, Toulouse, France. May 15-19, 2006.

[14] H. Bratt, L. Neumeyer, E. Shriberg, and H. Franco, "Collection and detailed transcription of a speech database for development of language learning technologies," Proc. ICSLP 98, pp. 1539-1542. Sydney, Australia. 1998.

[15] H. Franco, L. Neumeyer, Y. Kim, and O. Ronen, "Automatic pronunciation scoring for language instruction," Proc. ICASSP 97, pp. 1471-1474, Munich. 1997.

[16] J. L. Gauvain and C.H. Lee, "Maximum a posteriori estimation for multivariable Gaussian mixture observatioins of Markov models," IEEE Trans. on Speech and Audio Processing, 2(2), pp. 291-298, 1994.

[17] T. Joachims, "Making large-scale SVM learning practical," Advances in Kernel Methods - Support Vector Learning, B. Schölkopf, C. Burges, and A. Smola (ed.), MIT Press, 1999.

[18] S. Siegel and N. John Castellan, Jr., Nonparametric Statistics for the Behavioral Sciences, Second Edition. New York: McGraw-Hill, 1988.

[19] H. Franco, V. Abrash, K. Precoda, H. Bratt, R. Rao, and J. Butzberger, "The SRI EduSpeak(TM) system: Recognition and pronunciation scoring for language learning," Proc. InSTIL 2000, Dundee, Scotland, 2000.

[20] V. Digalakis, P. Monaco, P., and H. Murveit, Genones: Generalised mixture tying in continuous hidden Markov model-based speech recognizers, IEEE Trans. Speech and Audio Processing, vol. 4/4, pp. 281-289, 1996.