Maori Pronunciation Aid Speech Recognizer using HTK

Practical Work Report

Annie I-An Lu
ID: 1208065

The University of Auckland
Faculty of Engineering
Department of Electrical and Computer Engineering
Building 303, Room240
Level 2, Science Centre
38 Princes Street
Auckland

Phone: +62 9 373 7599 ext 88158
Fax: +64 9 373 7461
Email: ece-info@auckland.ac.nz

Work Period: 14/11/2011 – 10/02/2012
Report Date: 10/02/2012

Abstract

Four hundred hours of subprofessional electrical engineering work was completed under the supervision of Dr. Catherine Watson at the University of Auckland. This was done in response to pronunciation of the Maori language deteriorating due to influence from New Zealand English. A Maori pronunciation Aid, MPAi, was further developed on, to have a word recognizer that would give feedback on how well a word was pronounced. The HTK toolkit was used to create Hidden Markov Models (HMMs) sets for the word recognizer and the word recognition results of easily confusable Maori words were analyzed. Analysis showed that using both male and female training data produced a decent recognizer with a recognition result of 58% for tested young L1 male speakers and 52% for tested young L1 female speakers. It also showed that misrecognition occurred in the commonly confused vowels, proving that vowel sound differentiation has indeed been deteriorating. Live recognition test results on the proficient speaker showed that consonants were often misrecognised, likely due to a lack of data on the short consonant sounds when training the HMMs. The recognizer should be further developed before implementation into the MPAi program. Through the summer studentship, an introduction into phonetics, and signal processing in regards to speech recognition was gained. Appreciation of the important of having work experience, and an understanding of what doing a postgraduate course may be like was also acquired. This experienced aided me in deciding on the course of action to take after graduation.

Acknowledgements

I wish to extend my gratitude towards my supervisor, Dr. Catherine Watson, for giving me this learning opportunity, and for her constant support and guidance throughout this studentship period, and to the MAONZE group for providing the training and test Maori word recordings used in this project. I must also give a warm thank you to Peter Keegan for helping with the testing of the live audio output.

Table of Contents

Table of Figures

1.0 Introduction

As part of the required eight hundred hours of practical work experience required for the Bachelor of Engineering degree, four hundred hours of subprofessional electrical engineering work was undertaken between November 2011 and February 2012 at the University of Auckland Department of Electrical and Computer Engineering under the supervision of Dr. Catherine Watson. The summer studentship was to continue with the work on a Maori pronunciation aid, MPAi, to develop a word recognizer that would give the user feedback of how well a word was pronounced. This software was developed in response to the mispronunciation of common Maori words due to loss of Maori vowel sounds that are not present in New Zealand English. The HTK Toolkit for training Hidden Markov Models was used to create this word recognizer and the recognition results on commonly mispronounced Maori words was analyzed.

2.0 Structure of Organization

The University of Auckland- Faculty of Engineering's Department of Electrical and Computer Engineering (ECE) was founded in 1953 by Professor A.G. Bogle and two other academics. Starting off with a small number of students, the Department now supports a few hundred students per year due to successful programmes consisting of teaching specific technologies at the undergraduate level and shifting to having a strong focus on research at the postgraduate level. It maintains its high standards through industry partnerships and staff at the forefront of their respective fields [1]. Currently, the Department offers three undergraduate programmes, the Electrical and Electronics Engineering programme, Software Engineering programme, and the Computer Systems Engineering programme.

2.1 Layout (building/plant layout)

The ECE Department is situated on the first, second, and third floors of the university's city campus Science Building as seen in Figure 1 on the next page. There are many general and specialist laboratories for the student, researchers, and staff's use. The laboratory used in this summer project is the Biometrics Laboratory located on the second floor.

2.2 Amenities

Aside from laboratories with the necessary equipment and technologies required for the specific specializations and technologies, undergraduate students are provided with individual accounts for school computer use, as well as lockers, and toilets. In addition to these, staff and postgraduate students are also given access to staff rooms with couches, tables, microwaves, hot water, tea, hot chocolate and coffee.

The campus also has many food stalls, cafes, ATM machines, libraries, as well as an optometry clinic, medical centre, bookshop, pharmacy, and recreation centre. Parking and postal services are also provided [4].

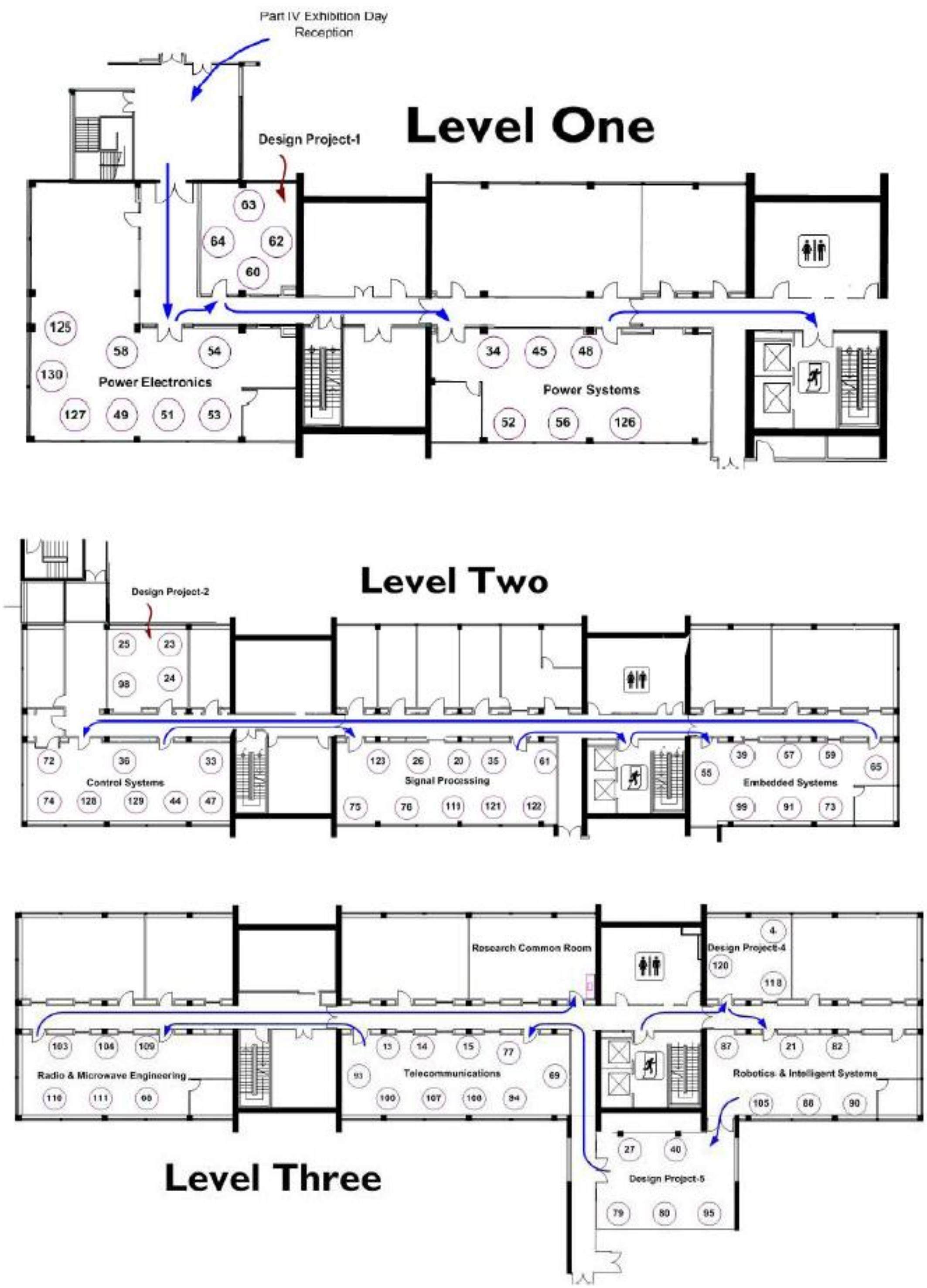

Figure 1: The ECE Department Layout in Building 303 on 38 Princes Street [2]

2.3 Staff (no of employees)

There are 58 staff members in the department [3]:

- 7 Professors
- 7 Associate Professors
- 18 Senior Lecturers
- 4 Lecturers
- 4 Senior Tutors
- 5 Postdoctoral Research Fellows
- 1 Professional Teaching Fellow
- 8 Technicians
- 4 Other Professional Staff (administration, office, departmental managers)

2.4 Health and Safety

Due to the number of students and staff using the specialist laboratories, and the nature of the work, there are health and safety officers for each level of the building, a first aid box in each laboratory and many on site fire wardens and first aiders [5]. Guidelines for laboratory use include [6]:

- Not being allowed in the laboratory without an access card
- Not being allowed to modify or repair equipment unless specifically authorized
- Not being allowed to use equipment unless specifically authorized and taught how to
- Not being allowed to smoke
- Not being allowed to eat
- Must only drink water from 'sipper' bottles
- Must wear suitable clothing, including covered shoes in all laboratories
- Work area must be kept clean and tidy
- Cables must be routed unobtrusively
- Equipment should be switched off after use
- Being responsible for contacting the relevant staff if any hazard or fault occurs

2.5 Facilities

The department's general undergraduate laboratories boasts more than a hundred computers and other state-of-the-art equipment such as digital oscilloscopes, function generators, multi-meters, power supplies, drills, and soldering stations.

The biometrics laboratory has an acoustic reflectometer for measuring the vocal tract shape, a sound-proof booth for reducing noise during sample recordings, and a variety of sophisticated recording equipment including headphones and microphones that were used for this research project.

3.0 Work Description

The project undertaken was to further develop on a Maori Pronunciation Aid, MPAi, which had previously been created to preserve Maori language pronunciation. The aim was to create a Maori word recognizer using HTK that would give a score of how well a word was pronounced once recognized.

3.1Motivations

Despite the Maori language being one of the official languages of New Zealand, the pronunciation of Maori has changed over the past century due to the country predominantly speaking English. As a result, vowel sounds in Maori have paralleled changes in New Zealand English; phonemes that do not occur in New Zealand English have led to a loss of vowel sounds that are in a similar vowel space in the Maori language [7].

Being one of the official languages of New Zealand, and under revitalization, it is important to ensure the preservation of the Maori language. However, as Maori is an indigenous language only used in New Zealand, with low demand elsewhere, there is currently no pronunciation aid that caters to Maori language teachers, Maori speakers, and people wishing to learn Maori.

To tackle this issue, a software application in the developmental stages, MPAi, has been created by the University of Auckland.

3.2Previous Work

A vowel pronunciation feedback program, Formant Aid, had previously been developed at the university, but proved difficult to understand for people unfamiliar with phonetics. As a result, a second project, MPAi, was started by Daniel Rivers and Jacinth Gutla, and continued by Byron Hui, under the supervision of Dr Catherine Watson, to make a user friendly pronunciation aid for common Maori words [8].

Aside from the formant aid available in MPAi, on the next page in Figure 2, it is hoped that an aid be developed that would show a score of how well a word is pronounced, on the next page in Figure 3, by training up hidden markov models (HMMs) of common Maori words, using 'gold standard' older generation Maori speaker pronunciations, with the HTK toolkit.

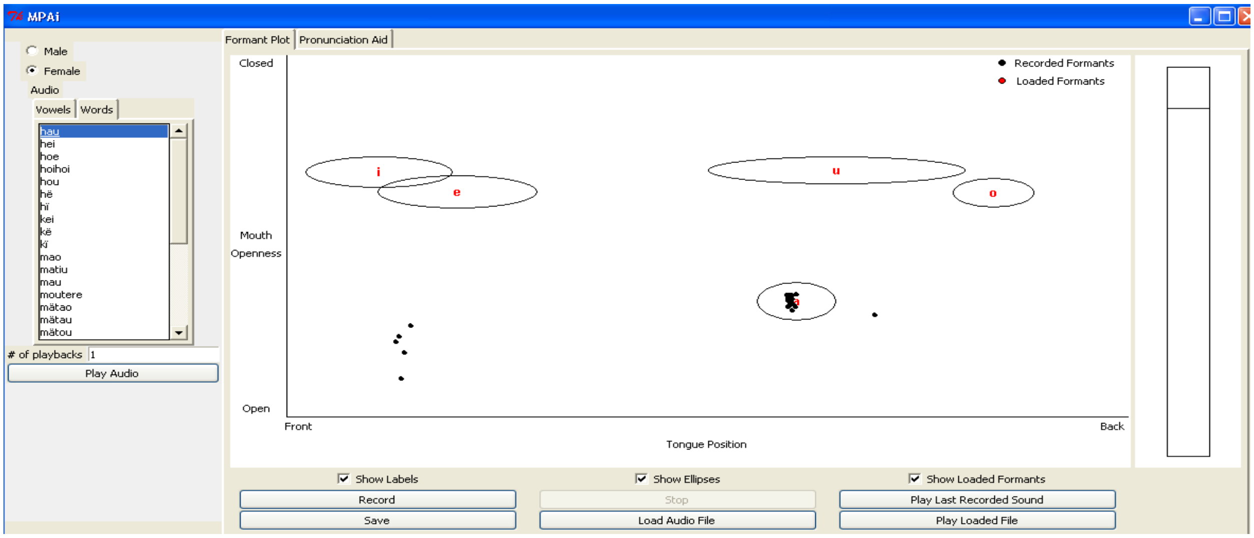

Figure 2: Format plot on MPAi

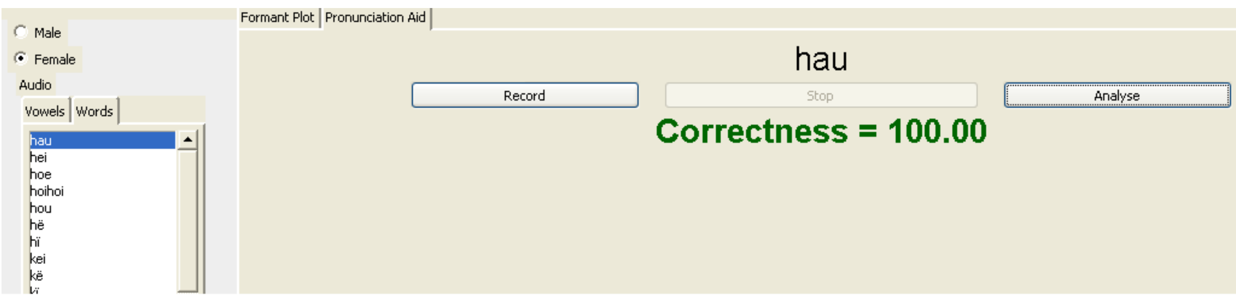

Figure 3: Pronunciation Score Aid on MPAi

3.3Hidden Markov Models

Phonemes, the smallest meaningful units of sound in a word, must be modeled statistically in a speech recognition system but are heavily influenced by their surrounding sounds, and so, changes from word to word. This variation is modeled with a three state Markov Chain; a starting state proceeded by the previous phoneme, a middle state, and an exiting state to the next phoneme. Each state has an associated output and is linked by transitions each with their own transition probably. By observing the output sequence of the word, one can infer the sequence of states that produced it.

A Hidden Markov Model (HMM), on the other hand, has more than one output per state; all outputs are possible at each state, and so, the output at each state also has its own output probability. The diagram below in Figure 4 illustrates this. States 1, 2, and 3 has state transitions to and away from it, shown by the arrows, with transitions probabilities a_{ii}, a_{jj}, a_{kk}, a_{ij}, a_{ji}, a_{jk}, a_{kj}, a_{ki}, and a_{ik}. Additionally, each state has outputs A, B, C, D, and E modelled by a probability distribution $b_1(s)$, $b_2(s)$, and $b_3(s)$. As each state can produce any of the outputs, the sequence of states cannot be retraced, and thus, is 'hidden' from the observer. In the case of speech recognition, the output probability density is defined over s, the spectrogram Mel frequency cepstral coefficients, or feature vectors, measured at each discrete time interval. To recognise speech, the most likely state sequence is found.

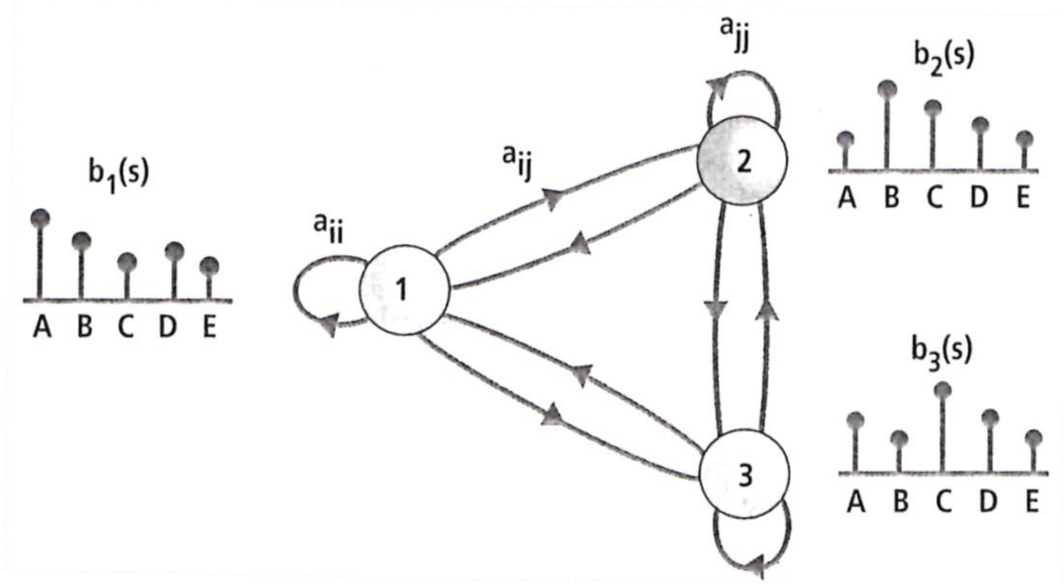

Figure 4: Diagram showing a three state HMM taken from [9]

3.4 HTK Toolkit

To model and manipulate the Maori language phoneme HMMs, the Hidden Markov Model Toolkit (HTK) was used. Developed by the Cambridge University Engineering Department, it is primarily used for the purpose of speech recognition [10]. Due to the wide range of support for speech recognition, including live audio recognition, it was chosen for this project.

3.5Maori Word Recognition HMM Training

To train the HMMs using HTK, several files had to be made. The files had to strictly adhere to the structure provided in the HTK Book [11, 12]. Some files had to be made manually while others just required commands to be typed into the command line; these commands have been included in Appendix A. All files made were stored in either the *user* directory, the *analysisOutput* directories or in the *HMMs* directories. The steps for creating a word recognizer were repeated several times for the different training data available; the main steps taken have been noted below.

1. A file of the grammar structures was first produced. As the aim for the recognizer was to do individual word recognition, this was unimportant. However, this step was still required to produce the *wordNetwork* file required in later steps.
   Files Produced: *grammar*, *wordNetwork*
   HTK Command used: HParse

2. A dictionary was created by first compiling a list of words to be recognized, and then creating a separate pronunciation dictionary file with the phonemes making up each word beside the word.
   Files Produced: *wordlist*, *dictionary*, *monophones0*, *monophones1, HDManFLOG, HDMANFLOGfullSet, HDMANFLOGOldFemale, HDManFLOGoldMale*
   HTK Command used: HDMan

3. Phone-level and word-level transcriptions of the prerecorded data was prepared.
   Files Produced: *wordTranscript.mlf, trainingWordTranscriptOldFemale, trainingWordTranscriptOldMale.mlf, trainingWordTranscriptWholeSet.mlf, testingAnnieWors.mlf, testL1YoungFemaleWordTranscript.mlf, testL1YoungMaleWordTranscript.mlf, phoneTranscript0.mlf, phoneTranscript1.mlf, phoneTranscriptOldFemaleTrainin0.mlf, phoneTranscriptOldFemaleTraining1.mlf, phoneTranscriptOldMaleTraining0.mlf, phoneTranscriptOldMaleTraining1.mlf, phoneTranscriptTrainingWholeSet0.mlf,phoneTranscriptTrainingWholeSet1.mlf, mkphones0EditScript4HLEd.led,mkphones1EditScript4HLEd.led*
   HTK Command used: HLEd

4. The script files and configuration file needed for training and testing were created. The configuration file determines how the signal processing is done, the parameters that are set, and thus, the resultant Mel frequency cepstral coefficient feature vectors produced. Additional configuration files were produced to suit later steps. Thorough understanding of the parameters available and what should have been set up was key to deciding the results of the training process. Configuration parameters of interest, and relevant concepts, can be seen in Appendix B, with the actual configurations parameters set, in Appendix C.
   Files Produced: *codeTestAnnie.scp, codeTestL1YoungFemale.scp, codeTestL1YoungMale.scp, codeTrain.scp, codeTrainL1YoungMale.scp, testAnnie.scp, testL1YoungMale.scp, train.scp, codeTrainOldFemale.scp, codeTrainOldMale.scp, testAnnie.scp, testL1YoungFemale.scp, testL1YoungMale.scp, train.scp, trainL1YoungMale.scp, trainOldFemale.scp, trainOldMale.scp, config0*
   HTK Command used: HCopy

5. An initial *flat start monophone* HMM model was created and then re-estimated.
   Files Produced: *hmm0, hmm0/hmmdefs, hmm0/macros, config1, hmm1, hmm1/hmmdefs, hmm1/macros, hmm2, hmm2/hmmdefs, hmm2/macros, hmm3, hmm3/hmmdefs, hmm3/macros*
   HTK Commands used: HCompV, HERest

6. The short pause, *sp* model, was then added in before further re-estimation.
   Files Produced: *sil.hed, hmm4, hmm4/hmmdefs, hmm4/macros, hmm5, hmm5/hmmdefs, hmm5/macros, hmm6, hmm6/hmmdefs, hmm6/macros*
   HTK Commands used: HHEd, HERest

7. Using forced alignment, the training data was realigned and re-estimated.
   Files Produced: aligned.mlf, alignedFullSet.mlf, alignedOldFemaleTraining.mlf, alignedOldMaleTraining.mlf, hmm7, hmm7/hmmdefs, hmm7/macros, hmm8, hmm8/hmmdefs, hmm8/macros, hmm9, hmm9/hmmdefs, hmm9/macros
   HTK Commands used: HVite, HERest

8. Triphone transcriptions were then made to improve recognition accuracy by linking three phonemes together, and thus reducing the error of confusing one sound with another. The resultant transcriptions were used for further HMM re-estimation
   Files Produced: mktri.hed, mktri.led, mktriFullSet.hed, mktriOldFemale.hed, mktriOldMale.hed, triphones1, triphoneTranscriptWINTRI.mlf, tirphoneTranscriptWINTRIfullSet.mlf, triphoneTranscriptWINTRIoldFemaleTrain.mlf, triphoneTranscriptWINTRIoldMaleTrain.mlf, hmm10, hmm10/hmmdefs, hmm10/macros, hmm11, hmm11/hmmdefs, hmm11/macros, hmm12, hmm12/hmmdefs, hmm12/macros
   HTK Commands used: HLEd, HHEd, HERest

9. Tied-state triphones and a phonetics rule file for clustering similar sounds were then created followed by final HMM re-estimation.
   Files Produced: lexicon.txt, triphoneDictionary, triphoneDictionaryFullSet, triphoneDictionaryOldFemale, triphoneDictionaryOldMale, fullPhoneListFullSet, fullPhoneListFullSet1, fullPhoneListFullSet2, fullPhoneListOldFemale, fullPhoneListOldFemale1, fullPhoneListOldFemale2, fullPhoneListOldMale, fullPhoneListOldMale1, fullPhoneListOldMale2, fullPhoneListOriginal, fullPhoneListOriginal1, fullPhoneListOriginal2, tree.hed, hmm13, hmm13/hmmdefs, hmm13/macros, hmm14, hmm14/hmmdefs, hmm14/macros, hmm15, hmm15/hmmdefs, hmm15/macros
   HTK Commands used: HDMan, HHEd, HERest

10. While some files used pre-recorded data which could be read from script files, for ease of use in real life, live recognition was made possible by not feeding in a script file. The final command decided upon for running the recognizer live allowed a file of the recognized transcriptions to be saved and played back the spoken word after each recording.
    HTK Commands used: HVite

11. The data recorded could be analyzed with analysis output files created and stored.
    Files Produced: analysisOutputFullSet, analysisOutputOldFemale, analysisOutputOldMale, analysisOutputOriginal, HResultsOut0TrainFullSet, HResultsOut1TestL1MaleFulleSet, HResultsOut2TestL1FemaleFullSet, HViteOut0TrainFullSet, HViteOut1TestL1MaleFullSet, HViteOut2TestL1FemaleFullSet, recout0aTrainFullSet,mlf, recout0bTrainFullSet.mlf, recout1aTestL1MaleFullSet.mlf, recout1bTestL1MaleFullSet.mlf, recout2aTestL1FemaleFullSet.mlf, recout2bTestL1FemaleFullSet.mlf, HResultsOut0TrainOldFemale, HResultsOut1TestL1MaleOldFemale, HResultsOut2TestL1FemaleOldFemale, HViteOut0TrainOldFemale, HViteOut1TestL1MaleOldFemale, HViteOut2TestL1FemaleOldFemale, recout0aTrainOldFemale.mlf, recout0bTrainOldFemale.mlf,recout1aTestL1MaleOldFemale.mlf, recout1bTestL1MaleOldFemale.mlf, recout2aTestL1FemaleOldFemale.mlf, recout2bTestL1FemaleOldFemale.mlf, HResultsOut0TrainOldMale, HResultsOut1TestL1MaleOldMale, HResultsOut2TestL1FemaleOldMale, HViteOut0TrainOldMale, HViteOut1TestL1MaleOldMale, HviteOut2TestL1FemaleOldMale, recout01TrainOldMale.mlf, recout0bTrainOldMale.mlf, recout1aTestL1MaleOldMale.mlf, recout1bTestL1MaleOldMale.mlf, recout2aTestL1FemaleOldMale.mlf, recout2bTestL1FemaleOldMale.mlf, HResultsOut0Train, HResultsOut1TestL1Male, HResultsOut2TestL1Female, HResultsOutAnnie, HviteOut0Train, HViteOut1TestL1Male, HViteOut2TestL1Female, HViteOutAnnie, recout01Train.mlf, recout0bTrain.mlf, recout1aTestL1Male.mlf, recout1bTestL1Male.mlf, recout2aTestL1Female.mlf, recout2bTestL1Female.mlf, recoutAnnieA.mlf, recoutAnnieB.mlf, trialingRecout.mlf
    HTK Commands used: HVite, HResults

There were many HMM versions trained with different sets of training data; each of these had their own files corresponding to each step in the training process. The files created, along with where they can be found, are described in Appendix D. In particular, it is important to note the batch files *go_v3_annie_part1, go_v3_annie_part2, go_v3_annie_part3*, and *analyse_v3_annie* were created to streamline the process of making HMMs and analysis data outputs. Another important file created was the *tree.hed* file; this is a phonetics rule file specifically for the Maori language. It clusters together similar sounds in the Maori language, in groups like consonants, vowels, voiced, unvoiced, fricative, nasal, alveolar, forward, central and back tongue positions, as well as monophthongs, and diphthongs. This was so that the recognizer was trained to take into account the similarities in each of the sound groups to improve its recognition ability. The phonetic rules used in this are listed in Appendix E.

As this was the first attempt at using triphones for Maori word recognition, many problems were initially faced with making sure the triphones were tied in the right way. If the phones were tied incorrectly, recognition rate would decrease instead of increase, as phones were associated incorrectly. An example of this was the word '*mau*' being recognized as '*mtau*', thus being incorrect. This was because instead of leaving the '*au*' phone by itself, or associating it as '*m+au'*, the default was to use the tied triphone '*t+au'* with the monophone '*m*' added to the front.

It was also the first attempt at live audio word recognition for the Maori language. Various parameters had to be included, or excluded, for this step, and many possible configuration files were tried before settling for *config6*. Apart from the configuration files, different commands were tried to find the one that worked best; the final command used, as well as possible commands to output data of the resulting recognition results have been included in Appendix A. The live recognition gave very poor results as the main test subject used during the development process could not speak Maori at all and thus it was initially hard to tell if the recognition was working until a proficient Maori speaker was asked to test the recognition. These results as well as results from the training and testing data sets, has been provided in Appendix F and will be discussed in the following section.

4.0 Word Recognition Evaluation

Three sets of HMMs were trained up, one using only the older generation male recordings (K group) as training data, one using only the older generation female recordings (R group), and one using both the older generation male and female recordings (K and R group). These were first tested with the data they were trained with, before being tested with younger generation Maori first language (L1) speaker recordings.

Results on the performance of each of the HMM sets made, based on percentage correctness as output by HTK, can be seen below in Table 1. The decision to train three sets of HMMs was to see if it would be possible to have drastically improved recognition results if male HMMs and female HMMs were trained separately. If this was the case, a female MPAi user would click to use the female HMMs and a male user would use the male HMMs. Testing results on younger generation L1 speakers showed that there was a slight improvement in female word recognition although there was a small drop in male word recognition from 61% to 58% for the HMM set trained using both K and R data. This small drop was deemed insignificant compared to the complications of using more than one set of HMM.

Table 1: Test Results for each of the different HMM sets made

Training Data Used	Performance on Training Data	Performance on Younger Generation L1 Male Speakers	Performance on Younger Generation L1 Female Speakers
Old Males (K)	85.71%	60.50%	34.44%
Old Females (R)	82.30%	38.66%	51.11%
Both (KR)	76.89%	57.98%	51.67%

Testing the HMMs trained using only the older generation male recordings with the younger generation female recordings gave a very low recognition result of 34% as did testing the older generation female HMMs with the younger generation males, which gave 39%. Such poor results were expected and no further analyses of them were carried out. It was decided that for all later stages, the HMM set with both the male and female training data would be used. Live audio recognition was also tried, and the results were analyzed to compare with prerecorded testing data. The following sections will discuss both prerecorded and live recognition results.

4.1 Female Voice's Impact on Recognition Results

From all the training and testing results of each HMM set, it is obvious that female word recognition gave lower results. This was expected as females have higher voices and therefore have formant frequencies in the wider bandwidths of higher frequency channels, making them more difficult to resolve than male voices.

Another reason for the poorer female results is due to females leading change in the Maori language; their pronunciation is less conservative and so their vowel merges, as mentioned earlier, are more pronounced than the males [7]. As the word list was compiled with commonly confused or mispronounced Maori words, it is expected for the females to perform worse given this list.

4.2 Likely Causes of Recognition Error

Consistent with the fact that the words provided were often mispronounced, many of the vowel substitution errors that occurred were in words such as *pao, pau, pou*, and *pö*; or in *tae*, and *tai*; or in *kë*, and *kei*; or *hoihoi* and *hoe*, where the vowel pronunciations were similar and easily confusable. Another error in vowel substitutions would be in the case where only the length of the vowel was different, and not the sound, like *kë* and *ke*; or *kï* and *ki*.

Apart from vowel substitutions, the recognizer was also prone to consonant substitutions. One reason would be that there consonants are very short in length and so there would be very little data on them, making them harder to recognize. Another reason would be that they are very similar. One example of this where consonants were often misrecognised was in the words *pai* and *tai*. Both 'p' and 't' are plosives, or oral stops, where airflow ceases and there would be no data. Other examples were *hau*, *mau, pau*, and *mät<u>au</u>*; *hou, pou* and *mät<u>ou</u>*; *mao, pao* and *mät<u>ao</u>*; *pae* and *tae*; *kë* and *hë, kï* and *hï; toi* and *hoihoi; toetoe* and *hoe*; or even *pö* and *tü* as both start with a plosive and end with a long vowel in a similar position in the mouth.

The number of consonant substitution errors was about the same as the number of vowel substitutions in all tests except for the live recognition test of the proficient Maori speaker. While the non-Maori speaker and phonetics expert had equal amounts of vowel and consonant substitutions errors, the proficient Maori speaker, who was asked to speak in a conservative manner, showed results with a much lower number for vowel substitutions as compared to consonant substitutions. This suggested that vowel substitution errors were caused by speaker mispronunciation while consonant substitution errors were caused by a lack of data provided to train and test the HMMs.

Apart from the substitution errors, there were also insertion and deletion errors. These occurred in words such as *tae* and *täne*. An extra 'n' sound was misrecognised in the word *tae* resulting in an insertion error, or the 'n' sound was missed out resulting in the word *täne* being misrecognised as *tae*. These occurred likely because the 'n' sound was short and easily added if the 'e' sound in tae was emphasized or drawn out, or if the word was spoken too quickly and the 'n' was missed. Insertion errors also occurred where the recognition was not sure of a particular vowel sound. For example, *pao*, may have been recognized as *pao pou* as the recognizer was unsure whether the sound was *pao,* or *pou*. Another source of insertion error could have been emphasis on the starting and ending sounds, resulting in bursts of energy observed as short extra words like k*ë, kï, kei, hë, hï, hei, pao, pau, pou, pai, tai, mai; toi,* or *tü.* Breath inhalation and exhalation could have caused sharp spikes of sound which may also have been misrecognised as an extra word. This would have been especially true in live recognition. Timing issues between button pressing and speaker talking during live recognition could also have resulted in the start or end of the words being missed in the recording, thus resulting in deletion errors.

4.3 Impact of Running Recognizer Live as Opposed to having Pre-Recorded Data

The training and testing data were not recorded in a recording booth and thus was non-ideal. Background noise in these pre-recorded data included conversation and infants crying. These would have affected the quality of the data used to train and test the HMMs.

This was worsened by the different sort of background noises during live recognition resulting in live recognition results being, as predicted, noticeably worse. Air conditioning and construction noise in the laboratory during live recognition, microphone quality, as well as the test subjects being female and unable to speak Maori resulted in nearly zero percent recognition. As the predominantly used test subject was a relatively high pitched female who could not speak Maori at all, and the older generation female training data had females with relatively low voices, testing the live recognition in the noisy laboratory, or in the isolated recording booth did not make any noticeable difference. Neither did using a broadcasting microphone that would have removed breath inhalation and exhalations, or other sharp spikes of sound. If this had been tested on the proficient male Maori speaker, there may have been noticeable difference.

5.0 Recommendations for Future Work

For the proficient Maori speaker, consonant substitution was a lot worse; as consonants tend to be of a high frequency, adjusting the lower and upper cutoff frequencies in the configuration parameters may result in consonants being ignored and thus would not affect the recognition results. Other configuration parameters could also be changed. This could be done to change from using linear prediction analysis to filterbank analysis to see which gives the better training result. Even if one chooses to continue using linear prediction analysis, the number of channels and filter order could be changed. Vocal tract length normalization via frequency warping could also be experimented with. If the parameters are set so a recognizer of usable quality is achieved, it could be implemented into MPAi to show a percentage correctness score, and the recognized word(s).

A completely different HMM training toolkit could also be used, instead of HTK, if it is deemed that HTK is unable to provide desirable results, or a completely different method of speech recognition such as using artificial neural networks could be tried. However, no matter what method is used, more and better quality training and testing data is required to get a desirable result. Once a desirable Maori word recognizer has been developed, it should be implemented into MPAi to achieve a finished product.

6.0 Conclusions

The summer studentship at the University of Auckland allowed me to learn about and develop an interest in phonetics, speech signal processing and its uses in the industry. In particular, I learnt about speech recognition and the problems faced during the revival of indigenous languages such as the Maori language. I learnt that Maori vowel merges occurring in younger generations does affect recognition results, and consonants being short means they require more data for a well trained recognizer to be made.

This experience not only allowed me to gain four hundred hours of the required eight hundred hours of practical work I need for my degree, it gave me a taste of what post graduate work may be like, aiding me in my decision of whether or not to go straight into the workforce after graduation, or to continue with my studies. I also gained an impression of what working in the workforce, with responsibilities, and tasks I must complete would be like. I realized the knowledge gained in class would help me but a lot of learning to do with the particular task is also needed, showing me the kind of continued learning I would need to do even after I enter the workforce. Overall, this has been a fruitful and wonderful experience.

References

[1] *Department of Electrical and Computer Engineering - About Us*. University of Auckland. Retrieved February 19, 2012 from: http://web.ece.auckland.ac.nz/uoa/home/about/our-department/about-us

[2] *Part 4 Exhibition Day 2010*. Faculty of Engineering, University of Auckland. Retrieved February 19, 2012 from: http://web.ece.auckland.ac.nz/webdav/site/ece/shared/about/events/documents/ECE%20exhibition%20for%20web.pdf

[3] *Department of Electrical and Computer Engineering – Locate a Staff*. University of Auckland. Retrieved February 19, 2012 from: http://web.ece.auckland.ac.nz/uoa/staff-directory

[4] *Facilities and Amenities*. University of Auckland. Retrieved February 19, 2012 from: http://www. auckland.ac.nz/uoa/facilities-amenities

[5] *Department of Electrical and Computer Engineering – Health and Safety*. University of Auckland. Retrieved February 19, 2012 from: http://web.ece.auckland.ac.nz/uoa/home/about/our-department/health-and-safety

[6] Department of Electrical and Computer Engineering – Health and Safety Manual. Faculty of Engineering, University of Auckland. Retrieved February 19, 2012 from: http://web.ece. auckland.ac.nz/webdav/site/ece/shared/about/our-dept/documents/ece-safety-manual.pdf

[7] Harlow, R., Keegan, P., King, J., Maclagan, M. and Watson, C. (2009) *Changing Pronunciation of the Maori Language: Implications for Revitalization.* Reyhner J, and Lockhead, L (Eds) Indigenous Language Revitalization Encouragement, Guidance & Lessons Learned Northern Arizona University, pp 75-86.

[8] Hui, B. (2011) Visual *Feedback for Language Production*. The University of Auckland.

[9] Comerford, R., Makhoul, J. and Schwartz,R. (1997) *The Voice of the Computer is Heard in the Land (and it Listens too!) [speech recognition]*. IEEE Spectrum, **34** (12), pp 39-47.

[10] *HTK Speech Recognition Toolkit*. Speech Vision and Robotics Group of the Cambridge University Engineering Department. Retrieved February 19, 2012 from: http://htk.eng.cam.ac.uk/

[11] Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X.A., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., Woodland, P. and Young, S. (2006) *The HTK Book (for HTK Version 3.4)*, Cambridge University Engineering Department, Cambridge.

[12] Tutorial: Create Acoustic Model – Manually. VoxForge. Retrieved February 19, 2012 from: http://www.voxforge.org/home/dev/acousticmodels/linux/create/htkjulius/tutorial

Appendix A: HTK Commands Used

This section will list the HTK commands used to train the Maori word recognizer; commands used for perl scripts from the HTK tutorial have also been included. It does not include information on how specific files that were manually made or modified.

HParse user/gramer user/wordNetwork

HLEd -l * -d user/dictionary -i user/phoneTranscipt0.mlf user/mkphones0EditScript4HLEd.led user/wordTranscript.mlf

HLEd -l * -d user/dictionary -i user/phoneTranscript1.mlf user/mkphones1EditScript4HLEd.led user/wordTranscript.mlf

HCopy -T 1 -C user/config0 -S user/codeTrain.scp

HCopy -T 1 -C user/config0 -S user/codeTestL1YoungMale.scp

HCompV -C user/config1 -f 0.01 -m -S user/train.scp -M hmm0 user/proto

At this point, the hmmdefs and macros file for hmm0, the first hidden markov model, are prepared.

HERest -C user/config1 -I user/phoneTranscript0.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm0/macros -H hmm0/hmmdefs -M hmm1 user/monophones0

HERest -C user/config1 -I user/phoneTranscript0.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm1/macros -H hmm1/hmmdefs -M hmm2 user/monophones0

HERest -C user/config1 -I user/phoneTranscript0.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm2/macros -H hmm2/hmmdefs -M hmm3 user/monophones0

The macros and hmmdefs files in hmm3 are copied to hmm4 and modified.

HHEd -H hmm4/macros -H hmm4/hmmdefs -M hmm5 user/sil.hed user/monophones1

HERest -C user/config1 -I user/phoneTranscript0.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm5/macros -H hmm5/hmmdefs -M hmm6 user/monophones1

HERest -C user/config1 -I user/phoneTranscript0.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm6/macros -H hmm6/hmmdefs -M hmm7 user/monophones1

HVite -l * -o SWT -b silence -C user/config1 -a -H hmm7/macros -H hmm7/hmmdefs -i user/aligned.mlf -m -t 250.0 150.0 1000.0 -y lab -I user/wordTranscript.mlf -S user/train.scp user/dictionary user/monophones1

HERest -C user/config1 -I user/aligned.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm7/macros -H hmm7/hmmdefs -M hmm8 user/monophones1

HERest -C user/config1 -I user/aligned.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm8/macros -H hmm8/hmmdefs -M hmm9 user/monophones1

HLEd -n user/triphones1 -l * -i user/triphoneTranscriptWINTRI.mlf user/mktri.led user/aligned.mlf

perl maketrihed user/monophones1 user/triphones1

HHEd -H hmm9/macros -H hmm9/hmmdefs -M hmm10 user/mktri.hed user/monophones1

HERest -C user/config1 -I user/triphoneTranscriptWINTRI.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm10/macros -H hmm10/hmmdefs -M hmm11 user/triphones1

HERest -C user/config1 -I user/triphoneTranscriptWINTRI.mlf -t 250.0 150.0 1000.0 -s user/stats -S user/train.scp -H hmm11/macros -H hmm11/hmmdefs -M hmm12 user/triphones1

HDMan -b sp -n user/fullPhoneList -g user/global.ded -l user/HDManFLOG user/triphoneDictionary user/lexicon.txt

The fullPhoneList file is then modified and a new file, fullPhoneList1, is created.

perl fixFullList.pl user/fullPhoneList1 user/fullPhoneList2

HHEd -H hmm12/macros -H hmm12/hmmdefs -M hmm13 user/tree.hed user/triphones1

HERest -C user/config1 -I user/triphoneTranscriptWINTRI.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm13/macros -H hmm13/hmmdefs -M hmm14 user/tiedList

HERest -C user/config1 -I user/triphoneTranscriptWINTRI.mlf -t 250.0 150.0 1000.0 -S user/train.scp -H hmm14/macros -H hmm14/hmmdefs -M hmm15 user/tiedList

For live recognition, the command used was:

HVite -H HMMsFullSet/hmm15/macros -H HMMsFullSet/hmm15/hmmdefs -C user/config6 -w user/wordNetwork -e -i output -g -p 0.0 -s 5.0 user/dictionary user/tiedList

There are many possible commands available for data analysis depending on what one wants to view. The final ones used were:

```
HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/train.scp -l * -T 4 -i analysisOutput/recout0aTrain.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList>analysisOutput/HViteOut0Train

HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/train.scp -l * -a -f -i analysisOutput/recout0bTrain.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList

HResults -t -I user/wordTranscript.mlf user/tiedList analysisOutput/recout0aTrain.mlf>analysisOutput/HResultsOut0Train

HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/testL1YoungMale.scp -l * -T 4 -i analysisOutput/recout1aTestL1Male.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList >analysisOutput/HViteOut1TestL1Male

HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/testL1YoungMale.scp -l * -a -f -i analysisOutput/recout1bTestL1Male.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList

HResults -t -I user/testL1YoungMaleWordTranscript.mlf user/tiedList analysisOutput/recout1aTestL1Male.mlf >analysisOutput/HResultsOut1TestL1Male

HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/testL1YoungFemale.scp -l * -T 4 -i analysisOutput/recout2aTestL1Female.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList >analysisOutput/HViteOut2TestL1Female

HVite -H -C user/config1 HMMs/hmm15/macros -H HMMs/hmm15/hmmdefs -S user/testL1YoungFemale.scp -l * -a -f -i analysisOutput/recout2bTestL1Female.mlf -w user/wordNetwork -p 0.0 -s 5.0 user/dictionary user/tiedList

HResults -t -I user/testL1YoungFemaleWordTranscript.mlf user/tiedList analysisOutput/recout2aTestL1Female.mlf >analysisOutput/HResultsOut2TestL1Female
```

Other possible commands tried that may be useful were:

```
HVite -H annieHMM/hmm15/macros -H annieHMM/hmm15/hmmdefs -S annietut/train.scp -C annietut/config1 -r 5.0 -l * -i annietut/recoutMon4.mlf -a -f -p 0.0 –s 5.0 annietut/dict annietut/tiedlist
```

An example output of this is:

```
#!MLF!#
"*/K004M_30.rec"
0 1300000 m[2] -967.352173 matiu
```

```
1300000 1600000 m[3] -265.874084
1600000 1700000 m[4] -91.405136
1700000 1900000 a[2] -148.113342
1900000 2400000 a[3] -332.721832
2400000 2800000 a[4] -291.498627
2800000 2900000 t[2] -98.029846
2900000 3000000 t[3] -113.751755
3000000 3100000 t[4] -126.341110
3100000 3200000 i[2] -68.932655
3200000 3400000 i[3] -130.870651
3400000 3700000 i[4] -177.051636
3700000 4000000 u[2] -191.565720
4000000 4400000 u[3] -253.272461
4400000 6000000 u[4] -905.351013
```

HVite -H annieHMM/hmm15/macros -H annieHMM/hmm15/hmmdefs -S annietut/train.scp -C annietut/config1 -r 5.0 -l * -i annietut/recoutMon5.mlf -a -p 0.0 -s 5.0 annietut/dict annietut/tiedlist

An example output of this is:

```
#!MLF!#
"*/K004M_30.rec"
0 6000000 matiu -4162.132324
```

HVite -H annieHMM/hmm15/macros -H annieHMM/hmm15/hmmdfs -S annietut/train.scp -C annietut/config1 -r 5.0 -l * -i annietut/recoutMon6.mlf -a -m -p 0.0 -s5.0 annietut/dict annietut/tiedlist

An example output of this is:

```
#!MLF!#
"*/K004M_30.rec"
0 1700000 m -1324.631348 matiu
1700000 2800000 a -772.333801
2800000 3100000 t -338.122711
3100000 3700000 i -376.854919
3700000 6000000 u -1350.083862
6000000 6000000 sp -0.105413
.
```

HVite -H annieHMM/hmm15/macros -H annieHMM/hmm15/hmmdfs -S annietut/train.scp -C annietut/config1 -r 5.0 -l * -i annietut/recoutMon6.mlf -a -m –b sil -p 0.0 -s5.0 annietut/dict annietut/tiedlist

Output example after adding a silence model, –b sil, is:

```
#!MLF!#
"*/K004M_30.rec"
0 900000 sil -548.633362 sil
900000 1700000 m -678.415283 matiu
1700000 2800000 a -772.333801
2800000 3100000 t -338.122711
3100000 3700000 i -376.854919
3700000 5700000 u -1159.414307
5700000 5700000 sp -0.105413
5700000 6000000 sil -190.418732 sil
.
```

HVite -H annieHMM/hmm15/macros -H annieHMM/hmm15/hmmdefs -S annietut/train.scp -C annietut/config1 -r 5.0 -o ST -l * -i annietut/recoutMon6.mlf -a -m -p 0.0 -s 5.0 annietut/dict annietut/tiedlist

```
#!MLF!#
"*/K004M_30.rec"
m matiu
a
t
i
u
sp
.
```

Appendix B: Configuration File Parameters available

This section gives an overview of configuration parameters of interest to do with the speech signal processing. The default configurations, where applicable, has been included.

GENERAL

SOURCEKIND = WAVEFORM (external source file to be converted into target kind target file)
TARGET KIND = MFCC_0_D_A (expects name of speech parameter type, delta and acceleration coefficients are to be calculated and added to static MFCC coefficients during coding) D and A are differences of the cepstral coefficients
Set to LPC to obtain filter parameters, and LPREFC to obtain reflection coefficients, LPCEPSTRA to generate linear prediction cepstra

SPEECH SIGNAL PROCESSING

SOURCERATE = rate at which source file is sampled. MUST BE SET MANUALLY WHEN USING DIRECT AUDIO INPUT.
TARGETRATE = 100000.0 (frame period 10ms)
WINDOWSIZE = 250000.0 #= 25ms (window size in units of 100ns) (segment of waveform used to determine each parameter vector, this is larger than the sample i.e. it is the frame)
ZMEANSOURCE = T– removes DC offset introduced during ADC
PREEMCOFF = 0.97 (pre-emphasis coefficient)
USEHAMMING=T (FFT uses Hamming window, gets rid of discontinuities at frame edges)
ADDITHER = q (positive value adds same amount of noise every time, negative value, the noise is random and changes between trials)

LINEAR PREDICTION ANALYSIS

LPCORDER = 12 (required filter order, here 12 coefficients per vector is produced)
NUMCEPS=12 (number of cepstra to generate)
CEPLIFTER = 22 scale higher order cepstral coefficient to larger values
NUMCHANS = 20 (number specifying number of filter bank channels used in analysis)

FILTERBANK ANALYSIS

USEPOWER = T (to set filterbank analysis to using power instead of magnitude for binning (binning = multiplying each FFT magnitude coefficient by their filter gain and summing up the result in each filterbank channel))
LOFREQ = 300 (set low frequency cut off and high frequency cut off for filterbank analysis)
HIFREQ = 3400

VOCAL TRACT LENGTH NORMALISATION

WARPFREQ = vocal tract length normalization filterbank analysis frequency axis warping factor. Factor<1. This is the amount that the frequency axis should be compressed to normalize. This is used in forced alignment.
WARPLCUTOFF – setting the cutoff-region for the warping function so that the HIFREQ and LOFREQ

boundaries are kept.
WARPUCUTOFF – upper cutoff limit

CEPSTRAL FEATURES
MFCC – target type for Mel Frequency Cepstral Coefficients. These are calculated from the log of the filterbank amplitudes using the Discrete Cosine Transform.
For this NUMCHANS (number of filterbank channels), NUMCEPTS (number of cepstral coefficients in linear prediction), CEPLIFTER need to be set.
This method allows the cepstral mean to be subtracted after analysis in the log cepstral domain so that long term spectral effects due to different microphones can be removed. This cepstral mean normalization is done by changing MFCC to MFCC_Z but cannot be done for live audio.

PERCEPTUAL LINEAR PREDICTION
This is an alternative to the Mel Frequency Cepstral Coefficients; it is based on the standard Mel frequency filterbank.

ENERGY MEASURES
MFCC_E_0 allows energy to be calculated; _0 includes the scaled 0-th cepstral coefficient.
ENORMALISE = T (Boolean, whether or not energy normalization is to be performed, F for live audio)
MUST BE SET MANUALLY TO FALSE
ESCALE – energy can be scaled by a factor
RAWENERGY = T allows energy to be calculated independent of windowing and pre-emphasis.

DELTA, ACCELERATION and THIRD DIFFERENTIAL COEFFICIENTS
The performance of the speech recognition system is improved by adding time derivatives to the static parameters.
_D is the first order regression coefficients; delta coefficients
_A is the second order regression coefficient, acceleration coefficients
_T is the third order regression coefficients

DIRECT AUDIO INPUT/OUTPUT
USESILDET =T Sets the input to be able to detect energy based speech/silence
SILENERGY Sets up the mean energy level for silence
MEASURESIL=T Causes detector to calibrate parameters from acoustic environment prior to sampling
SPEECHTHRESH level above which frames are classified as speech
SPCSEQCOUNT the frames that are counted as silence/ speech are grouped into windows consisting of a certain number (set by SPCSEQCOUNT) of consecutive frames.
SILGLCHCOUNT and **SPCGLCHCOUNT** if the number of frames marked as silence falls below a glitch count; the whole window is classed as speech.
SILMARGIN sets the number of frames before the actual detected start of speech to determine when to start processing the utterance.

Appendix C: Configuration Variables Set

Table C-1 shows the different configuration variables used during the training of Maori phoneme HMM. Config0 was the configuration file used during the initial stages of data preparation. Once training of HMMs started, config1 was used. Config6 was only used during live audio recognition.

Table C - 1: Configuration Variables used at different stages of HMM building

Config0	Config1	Config6
# Coding parameters TARGETKIND = MFCC_0 TARGETRATE = 100000.0 SAVECOMPRESSED = T SAVEWITHCRC = T WINDOWSIZE = 250000.0 USEHAMMING = T PREEMCOEF = 0.97 NUMCHANS = 26 CEPLIFTER = 22 NUMCEPS = 12 ENORMALISE = F SOURCEKIND = WAVEFORM SOURCEFORMAT = WAV TARGETFORMAT = HTK	# Coding parameters TARGETKIND = MFCC_0_D_A TARGETRATE = 100000.0 SAVECOMPRESSED = T SAVEWITHCRC = T WINDOWSIZE = 250000.0 USEHAMMING = T PREEMCOEF = 0.97 NUMCHANS = 26 CEPLIFTER = 22 NUMCEPS = 12 ENORMALISE = F	# Coding parameters TARGETKIND = MFCC_0_D_A TARGETRATE = 100000.0 SAVECOMPRESSED = T SAVEWITHCRC = T WINDOWSIZE = 250000.0 USEHAMMING = T PREEMCOEF = 0.97 NUMCHANS = 26 CEPLIFTER = 22 NUMCEPS = 12 ENORMALISE = F SOURCERATE = 625.0 SOURCEKIND = HAUDIO SOURCEFORMAT = HTK EUSESILDET = T MEASURESIL = F OUTSILWARN = T AUDIOSIG = -1 RECOUTPREFIX=audioTranscript

Appendix D: HTK File Descriptions

This section gives information about where each of the HMM sets and related files are stored.

The initial HMM set using older generation K speakers and younger generation L1 male speakers as training data is stored in the directory *HTKversions_all* under the directories:

HTK_creating
HTK_recreating_v11_trainedTogether_femaleTestDataAdded

The final three sets of HMMs trained were stored in the directory *HTKversions_important* under the directories:

HTK_recreating_v13_onlyWomanTrained	R training data
HTK_recreating_v14maleOnlyAdded	K training data
HTK_recreating_v15_fullSetAdded	KR training data

Although *HTK_recreating_v14maleOnlyAdded* and *HTK_recreating_v15_fullSetAdded* include the test results of R training HMM set, and *HTK_recreating_v14_fullSetAdded* includes the test results of K training HMM set, to use the HMM made from R training data, or to use the HMM set made from K training data, the correct directory's HMM set must be used. I.e. to use K trained HMM set, the HMMs to use are located in *HTK_recreating_v14maleOnlyAdded*

The HMM set used for live recognition, and that includes the live recognition results is stored in the directory *HTKversions_important* under the directory:

HTK_recreating_full_lastVersion

Under this directory, the following directories can be found:

analysisOutput Annie	Analysis output for Annie's HMM sets
Byron's HMMandUSER	Byron's HMM set made and User made
HMMsAnnie	HMM sets made by Annie
recordingRoomOutputTranscripts – Annie, Catherine, Peter	
test	testing data (Annie, Y, H)
train	training data (Y, K, R)
user	User directory made by Annie

analysisOutput Annie has all the training and testing data outputs from HTK. Under *analysisOutput Annie*, there are four files:

analysisOutputFullSet	HTK result output files for KR trained HMM set
analysisOutputOldFemale	HTK result output files for R trained HMM set
analysisOutputOldMale	HTK result output files for K trained HMM set
analysisOutputOriginal	HTK result output files for KY trained HMM set

Under each of these, files named like the following exist:

HResultsOut0TrainFullSet	HResult command output for training data (KR)
HResultsOut1TestL1MaleFullSet	HResult command output for Y
HResultsOut2TestL1FemaleFullSet	HResult command output for H
HViteOut0TrainFullSet	Phone recognition at each time interval for training
HViteOut1TestL1MaleFullSet	Phone recognition at each time interval for Y
HViteOut2TestL1FemaleFullSet	Phone recognition at each time interval for H
recout0aTrainFullSet.mlf	Phone recognition transcript of word for training
recout0bTrainFullSet.mlf	Phone recognition transcript at each time interval
recout1aTestL1MaleFullSet.mlf	Phone recognition transcript of word for Y
recout1bTestL1MaleFullSet.mlf	Phone recognition transcript at each time for Y
recout2aTestL1FemaleFullSet.mlf	Phone recognition transcript of word for H
recout2bTestL1FemaleFullSet.mlf	Phone recognition transcript at each time for H

HMMsAnnie contains four HMM sets made by Annie:

HMMsFullSet	HMM set trained using K and R
HMMsOldFemale	HMM set trained using R
HMMsOldMale	HMM set trained using K
HMMsOriginal	HMM set trained using K and Y

recordingRoomOutputTranscripts - Annie, Catherine, Peter contains .wav file recordings and output transcripts of Annie Lu, Catherine Watson and Peter Keegan during live recognition.

The *user* directory contains the most important files. These were the files made and used for training the HMM sets. *User* contains all the files used for training the HMM sets:

aligned.mlf	A monophone transcription made after forced alignment (KY)
alignedFullSet.mlf	A monophone transcription made after forced alignment (KR)
alignedOldFemaleTraining.mlf	A monophone transcription made after forced alignment (R)
alignedOldMaleTraining.mlf	A monophone transcription made after forced alignment (K)
codeTestAnnie.scp	Script made for running the test data in 'Annie'
codeTestL1YoungFemale.scp	Script made for running the test data H
codeTestL1YoungMale.scp	Script made for running the test data Y
codeTrain.scp	Script for running the training data KR
codeTrainL1YoungMale.scp	Script for running the training data Y
codeTrainOldFemale.scp	Script for running the training data R
codeTrainOldMale.scp	Script for running the training data K
config0	Configuration file used initially at the start of training
config1	This configuration file is changed part way through the training
config2	One of the config files made for live audio
config3	One of the config files made for live audio
config4	One of the config files made for live audio
config5	One of the config files made for live audio
config6	The final configuration file used for live audio
dictionary	List of Maori words and there monophone pronunciations
fullPhoneListFullSet	Full list of monophones for HMM set trained on KR
fullPhoneListFullSet1	Initial edit of full list of monophones for HMM set trained on KR
fullPhoneListFullSet2	Final edit of full list of monophones for HMM set trained on KR
fullPhoneListOldFemale	Full list of monophones for HMM set trained on R
fullPhoneListOldFemale1	Initial edit of full list of monophones for HMM set trained on R
fullPhoneListOldFemale2	Final edit of full list of monophones for HMM set trained on R
fullPhoneListOldMale	Full list of monophones for HMM set trained on K
fullPhoneListOldMale1	Initial edit of full list of monophones for HMM set trained on K
fullPhoneListOldMale2	Final edit of full list of monophones for HMM set trained on K
fullPhoneListOriginal	Full list of monophones for HMM set trained on KY
fullPhoneListOriginal1	Initial edit of full list of monophones for HMM set trained on KY
fullPhoneListOriginal2	Final edit of full list of monophones for HMM set trained on KY
global.ded	Edit script for HDMan command
grammer	Files containing grammar rules
HDManFLOG	Log output from HDMan command for HMM set training on KY
HDManFLOGfullSet	Log output from HDMan command for HMM set trained on KR
HDManFLOGOldFemale	Log output from HDMan command for HMM set trained on R
HDManFLOGoldMale	Log output from HDMan command for HMM set trained on K
lexicon.txt	A dictionary list

mkphones0EditScript4HLEd.led	First edit script for HLEd command
mkphones1EditScript4HLEd.led	Second edit script for HLEd command
mktri.hed	Edit script for HHEd command for HMM set trained on KY
mktri.led	Edit script for HLEd
mktriFullSet.hed	Edit script for HHEd command for HMM set trained on KR
mktriOldFemale.hed	Edit script for HHEd command for HMM set trained on R
mktriOldMale.hed	Edit script for HHEd command for HMM set trained on K
monophones0	List of monophones without *sp* (short pause) model
monophones1	List of monophones with the *sp* model added
phoneTranscript0.mlf	Phone transcription 1 made during the training using KY
phoneTranscript1.mlf	Phone transcription 2 made during the training using KY
phoneTranscriptOldFemaleTraining0.mlf	Phone transcription 1 made during training using R
phoneTranscriptOldFemaleTraining1.mlf	Phone transcription 2 made during training using R
phoneTranscriptOldMaleTraining0.mlf	Phone transcription 1 made during training using K
phoneTranscriptOldMaleTraining1.mlf	Phone transcription 2 made during training using K
phoneTranscriptTrainingWholeSet0.mlf	Phone transcription 1 made during training using KR
phoneTranscriptTrainingWholeSet1.mlf	Phone transcription 2 made during training using KR
proto	Prototype file used for making the *proto* file in hmm0
sil.hed	Script used for HHEd command
stats_fullSet	Stats data produced while making hmm12 for KR
stats_oldFemale	Stats data produced while making hmm12 for R
stats_oldMale	Stats data produced while making hmm12 for K
stats_original	Stats data produced while making hmm12 for KY
testAnnie.scp	Script with list of Annie test files
testingAnnieWords.mlf	Master label transcription file of Annie's test words
testL1YoungFemale.scp	Script with list of H test files
testL1YoungFemaleWordTranscript.mlf	Master label transcription file for test data H
testL1YoungMale.scp	Script with list of Y test files
testL1YoungMaleWordTranscript.mlf	Master label transcription file for test data Y
tiedList	List of tied triphones
train.scp	Script with list of training files for KY
trainingWordTranscriptOldFemale.mlf	Master label transcription file of R training data
trainingWordTranscriptOldMale.mlf	Master label transcription file of K training data
trainingWordTranscriptWholeSet.mlf	Master label transcription file of KR training data
trainL1YoungMale.scp	Script with list of Y training files
trainOldFemale.scp	Script with list of R training files
trainOldMale.scp	Script with list of K training files
tree.hed	Phonetics rule file for sounds in the Maori language

trees	An edited version of the *tree.hed* file
triphoneDictionary	Dictionary list made after triphones are made for KY
triphoneDictionaryFullSet	Dictionary list made after triphones are made for KR
triphoneDictionaryOldFemale	Dictionary list made after triphones are made for R
triphoneDictionaryOldMale	Dictionary list made after triphones are made for K
triphones1	Triphone list
triphoneTranscriptWINTRI.mlf	Training list triphone transcriptions for KY
triphoneTranscriptWINTRIfullSet.mlf	Training list triphone transcriptions for KR
triphoneTranscriptWINTRIoldFemaleTrain.mlf	Training list triphone transcriptions for R
triphoneTranscriptWINTRIoldMaleTrain.mlf	Training list triphone transcriptions for K
wordlist	List of Maori words
wordNetwork	Word Network formed
wordTranscript.mlf	Master label transcription file for KY

Appendix E: Phonetic Rules for Maori Sound Classification

The following shows the phonetic rules in the tree.hed file and only includes the parts of the file with Maori sound classification. The location to find the entire tree.hed file is found in Appendix D.

Silence { *+sil }

Stop { *+p,*+t,*+k }

Nasal { *+m,*+n,*+N }

Fricative { *+f, *+h }

Liquid { *+r,*+w }

Vowel
{ *+a,*+e,*+i,*+o,*+u,*+a:,*+e:,*+i:,*+o:,*+u:,*+ae,*+ai,*+ao,*+au,*+ea,*+ei,*+eo,*+eu,*+ia,*+ie,*+io,*+iu,*+oa,*+oe,*+oi,*+ou,*+ua,*+ue,*+ui,*+uo }

C-Front { *+p,*+m,*+f,*+w }

C-Central { *+t,*+n,*+r }

C-Back { *+k,*+N,*+h }

V-Front { *+i,*+i:,*+e,*+e: }

V-Central { *+u,*+u: }

V-Back { *+o,*+o:,*+a,*+a: }

Front { *+p,*+m,*+f,*+w,*+i:,*+e,*+e: }

Central { *+t,*+n,*+r,*+u,*+u: }

Back { *+k,*+N,*+h,*+o,*+o:,*+a,*+a: }

Fortis { *+p,*+t,*+k,*+f }

UnFortLenis { *+m,*+n,*+N,*+h,*+r,*+w }

Coronal { *+t,*+n,*+r }

NonCoronal { *+p,*+m,*+k,*+N,*+f,*+h,*+w }

Anterior { *+p,*+m,*+t,*+n,*+f,*+w }

NonAnterior { *+k,*+N,*+r }

Continuent { *+m,*+n,*+N,*+f,*+h,*+r,*+w }

NonContinuent { *+p,*+t,*+k }

NonStrident { *+f,*+h }

UnStrident { *+p,*+m,*+t,*+n,*+k,*+N,*+r,*+w }

Glide { *+h,*+r,*+w }

Unvoiced-Cons { *+p,*+t,*+k,*+f,*+h }

Voiced-Cons { *+m,*+n,*+N,*+r,*+w }

Unvoiced-All { *+p,*+t,*+k,*+f,*+h,*+sil }

Long
{ *+a:,*+e:,*+i:,*+o:,*+u:,*+ae,*+ai,*+ao,*+au,*+ea,*+ei,*+eo,*+eu,*+ia,*+ie,*+io,*+iu,*+oa,*+oe,*+oi,*+ou,*+ua,*+ue,*+ui,*+uo }

Short { *+a,*+e,*+i,*+o,*+u }

Dipthong
{ *+ae,*+ai,*+ao,*+au,*+ea,*+ei,*+eo,*+eu,*+ia,*+ie,*+io,*+iu,*+oa,*+oe,*+oi,*+ou,*+ua,*+ue,*+ui,*+uo }

Fronting { *+ae,*+ai,*+ei,*+oe,*+oi,*+ue,*+ui }

High { *+i,*+i:,*+u,*+u: }

Medium { *+e,*+e:,*+o,*+o:,*+m }

Low { *+a,*+a:,*+ao,*+au,*+eu }

Rounded { *+o,*+o:,*+u,*+u:,*+ou,*+w }

Unrounded { *+a,*+a:,*+e,*+e:,*+i,*+i:,*+m,*+h,*+r }

Unvoiced-Stop { *+p,*+t,*+k }

Appendix F: Training Results

This section presents how the HTK output results are displayed, and also shows how these output results were then analysed.

An example of the HTK result output is seen below. The word after *LAB:* shows the intended word, and the word after *REC:* shows the recognized word. The word recognition overall results were displayed as a percentage, with *%Corr* meaning percentage correctness, *ACC* meaning accuracy, *N* shows the number of words tested, *H* shows the number of words recognized correctly, *D* shows the number of deletion errors, *S* shows the number of substitution errors, and I shows the number of insertion errors. The accuracy percentage is lower as it takes insertion errors into account.

```
Aligned transcription: R008M_14.lab vs R008M_14.rec
 LAB: hë
 REC: kë
Aligned transcription: R008M_15.lab vs R008M_15.rec
 LAB:     hei
 REC: pae kï
Aligned transcription: R008M_19.lab vs R008M_19.rec
 LAB: mätao
 REC: mätau
Aligned transcription: R008M_26.lab vs R008M_26.rec
 LAB: mao
 REC: mao pou
Aligned transcription: R008M_28.lab vs R008M_28.rec
 LAB:     moutere
 REC: mau tënei
====================== HTK Results Analysis =======================
  Date: Thu Feb 02 13:38:00 2012
  Ref : user/trainingWordTranscriptWholeSet.mlf
  Rec : analysisOutputFullSet/recout0aTrainFullSet.mlf
------------------------ Overall Results --------------------------
SENT: %Correct=72.09 [H=390, S=151, N=541]
WORD: %Corr=76.89, Acc=61.74 [H=416, D=0, S=125, I=82, N=541]
===================================================================
```

Each output was then looked at to see exactly where the errors were occurring in each word. The following set of tables displays whether the consonants or vowels were misrecognised, and whether the misrecognised vowels were misrecognised as those that they are commonly confused with. In all the following tables, *K* stands for the old male generation used for training, *R* stands for the old female generation used for training, *Y* stands for the L1 young males used for testing, and *H* stands for the L1 young females used for testing.

These symbols were used:

- () random noise
- / can't decide which
- O true
- X false

Table F - 1: Misrecognition results on the Y Group using L and R training data

Intended Word	Recognized Word(s)	Misrecognised Consonant?	Misrecognised Vowel?	Confusable Vowel Pair Exists?	Misrecognised as the other vowel of the pair
tënei	kë (tü) toi	OO	XO	O	XX
tënei	kë (tai) hoe	OO	XO	O	XX
hau	Hei	X	O	O	X
hau	Hou	X	O	O	O
hou	pau	O	O	O	O
hou	Pou	O	X	O	X
hou	Hau	X	O	O	O
pao	Mao	O	X	O	X
pao	Pao (pao)	X	X	O	X
pau	Pou	X	O	O	O
pau	Pou	X	O	O	O
pou	Pau	X	O	O	O
pö	tü	O	O	O	X
pai	Pae	X	O	O	O
pai	Toi	O	O	O	X
pai	Hoe	O	O	O	X
kei	kï	X	O	O	X
kei	Tae (kï)	O	O	O	O
kï	kë	X	O	X	X
hë	Kë	O	X	O	X
hë	kï	O	O	O	X
hë	kï	O	X	O	X
hë	kë	O	X	O	X
hei	Kei	O	X	O	X
hei	kï	O	O	O	X
hei	kï	O	O	O	X
hei	Hoe	X	O	O	O
hï	Kï	O	X	X	X

hï	Kï	O	X	X	X
hï	Kï	O	X	X	X
hï	Kï (kï)	O	X	X	X
tae	Pae	O	X	O	X
tae	pai	O	O	O	O
tae	Pae	O	X	O	X
tai	Pai	O	X	O	X
tai	Hoe	O	O	O	X
mätau	Mätou	X	O	O	O
mätau	Mao pou (tai)	XO	O	O	O
mätou	Mätau	X	O	O	O
mätou	Mätau	X	O	O	O
mätou	Mätau	X	O	O	O
toetoe	Toi toi	X	O	O	O
toetoe	Toi toi	X	O	O	O
toetoe	Hoe (pao) hoe	OO	X	O	X
hoihoi	Hoe hoe	X	O	O	O
mau	Pau	O	X	O	X
moutere	Hou kë kë	OO	XOO	O	XOO
tü	kï	O	O	X	X
tü	kï	O	O	X	X
tü	kï	O	O	X	X
matiu	kï	O	O	X	X

Table F - 2: Misrecognition results on the H Group using K and R training data

Intended Word	Recognized Word(s)	Misrecognised Consonant?	Misrecognised Vowel?	Confusable Vowel Pair Exists?	Misrecognised as the other vowel of the pair
tënei	Kë tae	OO	XO	O	XO
tënei	Kë toi	OO	XO	O	XX
täne	Täne (kë)	X	X	X	X
täne	Täne (kë)	X	X	X	X
täne	Hau hë	O	OO	X	X
täne	Tae tai	XO	OO	X	X
täne	Hou	O	O	X	X
hau	Hau (tü)	X	X	O	X
hau	Tae	O	O	O	X
hou	Pou	O	X	O	X
hou	hoe	X	O	O	O
hou	Hei	X	O	O	X
pao	(Mätao) pou	X	O	O	O
pao	Pao (tae)	X	X	O	X

pao	Pao/pou	X	X	O	X
pau	Hou	O	O	O	O
pau	Hou	O	O	O	O
pau	Hau	O	X	O	X
Pau	hau	O	X	O	X
pau	Tae	O	O	O	X
pou	Hou	O	X	O	X
pou	tü	O	O	O	X
pou	Hou	O	X	O	X
pou	Hou	O	X	O	X
pou	Tae	O	O	O	X
pö	Pö/pou	X	X/O	O	X/O
Pö	Hou	O	O	O	O
pö	Hou	O	O	O	O
pö	Hau/pao	O/X	O/O	O	O/X
pö	Pao	X	O	O	X
pai	Hoe	O	O	O	X
pai	Pai/ pou	X	X/O	O	X/X
pai	Pai/kei	X/O	X/O	O	X/X
pai	Pou	X	O	O	X
pai	Pae/tae	O/O	O/O	O	O/O
kë	Hë	X	X	O	X
kei	Hei	X	X	O	X
kei	Hei	X	X	O	X
kei	Hei	X	X	O	X
kï	Kë	X	O	X	X
kï	Kë	X	O	X	X
kï	kei	X	O	X	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hë	(Täne) kë	O	X	O	X
hë	Kë	O	X	O	X
hï	Kï	O	X	X	X
hï	Kë	O	O	X	X
hï	Kï	O	X	X	X
hï	Kï	O	X	X	X
hï	Kei	O	O	X	X
tae	Pae	O	X	O	X
tae	Pae	O	X	O	X
tae	Pae (kë)	O	X	O	X
tae	Pae	O	X	O	X
tae	Pae	O	X	O	X
tai	Tae (pou)	O	O	O	O
Tai	Pai	O	X	O	X
tai	Pae (kei)	O	O	O	O
tai	Pai	O	X	O	X

tai	Pae	O	O	O	O
mätao	Mätao (pae)	X	X	O	X
mätao	Mätao (pou)	X	X	O	X
mätao	Mätao (pou)	X	X	O	X
mätao	Mätao (pou)	X	X	O	X
mätau	Mätou	X	O	O	O
mätau	Mätou	X	O	O	O
mätou	Mätao (hoe)	X	O	O	O
mätou	Mätau	X	O	O	O
mätou	Mätao (hoe)	X	O	O	O
mätou	Mätau (tae)	X	O	O	O
Toetoe	Hoe hoe	O	X	O	X
toetoe	Toi hoe (pou)	XO	OX	O	OX
toetoe	Toi hoe (kë)	XO	OX	O	OX
toetoe	Toi toi	X	OO	O	OO
Toetoe	Hoe hoe	OO	XX	O	XX
toetoe	Hoe tënei	OOO	XO	O	XX
toi	Hoihoi	O	X	O	X
toi	Hoe (kei)	O	O	O	O
toi	Hoe (kei)	O	O	O	O
toi	Hoe (kei)	O	O	O	O
hoihoi	Hoe hoe	X	OO	O	OO
hoihoi	Hoihoi (kë)	X	X	O	X
hoihoi	Hoe hoe (kei)	X	OO	O	OO
hoe	Hoe (kei)	X	X	O	X
hoe	Hoihoi	X	O	O	O
hoe	Hoe/toi	X/O	X/O	O	X/O
mau	(Tü) mau (tü)	X	X	O	X
mau	Mau (tü)	X	X	O	X
mau	(Kë) pae	O	O	O	X
moutere	Mau	X	O	O	O
moutere	Mao/hou kë kë	X/O O O	O/X O O	O	O/X O O
moutere	Hou/mau tënei	O/X X O	X/O O O	O	X/O O O
moutere	Hoihoi tai kë	OXO	OOO	O	XXO
moutere	Tae tënei	OXO	OOO	O	XOO
Tü	Mau (kë kë)	O	O	X	X
Tü	Kë	O	O	X	X
Tü	Kei	O	O	X	X
Tü	Kei	O	O	X	X
matiu	(Tü) matiu (tae) (pae)	X	X	X	X
matiu	mätao (kë)	X	O	X	X

Table F - 3: Misrecognition results on the Y Group using K training data

Intended Word	Recognized Word(s)	Misrecognised Consonant?	Misrecognised Vowel?	Confusable Vowel Pair Exists?	Misrecognised as the other vowel of the pair
tënei	Kë hei	OO	X	O	X
tënei	(Kei) tënei	X	X	O	X
täne	Pau hoe	OO	OO	O	XX
hau	Pai	O	O	O	X
hou	Pau	O	O	O	O
hou	Pou	O	X	O	X
pao	Mao	O	X	O	X
pao	Mao	O	X	O	X
pao	Pao (pao)	X	X	O	X
pau	Mau	O	X	O	X
pai	Pae	X	O	O	O
pai	pae	X	O	O	O
kei	Kï	X	O	O	X
kei	Tae (kï)	O	O	O	X
kei	Kë	X	O	O	O
kï	Kë	X	O	X	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hei	kë	O	O	O	O
hei	Hï	X	O	O	X
hï	Kï	O	X	O	X
hï	Kï	O	X	O	X
hei	Hï	X	O	O	X
hï	Kï	O	X	X	X
hï	Kï	O	X	X	X
hï	(Kë) hï	X	X	X	X
tae	Pae	O	X	O	X
tae	Pae	O	X	O	X
tae	Pae	O	X	O	X
tai	Pai	O	X	O	X
tai	Pai	O	X	O	X
tai	Pai (hoe)	O	X	O	X
mätao	Mätou/pao	X/O	XO/X	O	XO/X
mätao	Mätao/hou	X/O	X/O	O	X/O
mätau	Mätau (tai)	X	X	O	X
mätau	Mätou (tai)	X	O	O	O
mätou	Mätau	X	O	O	O
mätou	Mätau/pou	X/O	XO/X	O	XO/X

toetoe	Toi toi	X	O	O	O
toetoe	Toi toi	X	O	O	O
toetoe	Toi toi	X	O	O	O
toetoe	Toi (pao) hoe	XO	OX	O	OX
hoe	Hoihoi	X	O	O	O
hoe	Toi	O	O	O	O
hoe	Hou/toi	X/O	O/O	O	X/O
moutere	Mau toi	X	OX	O	OX
moutere	Mau tënei	XXO	OOO	O	OOO
moutere	Hou kë kë	OOO	XOO	O	XOO
moutere	Mau kë	XO	OO	O	OO
tü	Kë	O	O	X	X
tü	Kï	O	O	X	X
tü	Kï	O	O	X	X
matiu	Toi kï	OO	OO	X	XX

Table F - 4: Misrecognition results on the H Group using R training data

Intended Word	Recognized Word(s)	Misrecognised Consonant?	Misrecognised Vowel?	Confusable Vowel Pair Exists?	Misrecognised as the other vowel of the pair
tënei	Kë tae	OO	XO	O	XO
tënei	Hë toi	OO	XO	O	XX
tënei	Kë/tae kë	O/X O	X/O O	O	X/X O
täne	Hau hë	OO	OO	O	OO
täne	Täne (kë)	X	X	O	X
täne	Täne (kë)	X	X	O	X
täne	Tai	X	O	O	X
täne	hë	O	O	O	O
Hau	Tü	O	O	O	X
Hau	Pau	O	X	O	X
hau	Pau	O	X	O	X
hau	Pau (kë)	O	X	O	X
hou	Hoe	X	O	O	O
hou	Tae (kë)	O	O	O	X
pao	Mao	O	X	O	X
pao	Mätao/pou	X/O	X/O	O	X/O
pao	Pao/pou	X	X/O	O	X/O
pau	Hau	O	X	O	X
pau	Hau	O	X	O	X
pau	Hau	O	X	O	X

pau	Hau	O	X	O	X
pau	Tae/tai	O/O	O/O	O	X/X
pou	Hou	O	X	O	X
pou	Hou	O	X	O	X
pou	Hou	O	X	O	X
pou	Hau	O	O	O	O
pou	Hau	O	O	O	O
pou	Pau	X	O	O	O
pou	Pau	X	O	O	O
pö	Hou	O	O	O	O
pö	Hou	O	O	O	O
pö	Hou	O	O	O	O
pö	hau	O	O	O	O
pö	hë	O	O	O	X
pö	Pao	X	O	O	X
pai	Tae	O	O	O	O
pai	Tae (kei)	O	O	O	O
pai	Tai	O	X	O	X
pai	Tai	O	X	O	X
pai	Pae	X	O	O	O
Pae	Hau	O	O	O	X
Pae	Täne	O	O	O	O
pae	Pae (kë)	X	X	O	X
pae	Tae	O	X	O	X
kei	Tai	O	O	O	X
kei	Hei	O	X	O	X
kei	Hei	O	X	O	X
kei	Kë	X	O	O	O
Kei	Hei	O	X	O	X
Kei	Kë	X	O	O	O
kï	Hou	O	O	X	X
kï	Kë	X	O	X	X
Kï	kë	X	O	X	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hë	Kë	O	X	O	X
hei	Kei	O	X	O	X
hei	Tae	O	O	O	O
hï	Kei	O	O	X	X
hï	Kï	O	X	X	X
hï	kei	O	O	X	X
hï	Kï	O	X	X	X
hï	kë	O	O	X	X
tae	(Pao) pae (kë)	O	X	O	X
tae	Pao/pai	O/O	O/O	O	X/O

tai	Hoe	O	X	O	O
tai	Pai	O	X	O	X
tai	Tae	X	O	O	O
mätao	Mätou	X	XO	O	XO
mätao	Mätao (pou)	X	X	O	X
mätou	Mätau (hë)	X	XO	O	XO
mätou	Mätau	X	XO	O	XO
toetoe	Hoe hoe	OO	XX	O	XX
toetoe	Toetoe (pou)	X	XX	O	XX
toetoe	Toi hoe (kë)	XO	OX	O	OX
toetoe	Hoe hoe	OO	XX	O	XX
Toetoe	Hoe hoe	OO	XX	O	XX
toetoe	Toi hoe	XO	OX	O	OX
Toi	Hoihoi	O	X	O	X
toi	Hoe (kei)	O	O	O	X
toi	Hoe	O	O	O	X
hoihoi	Hoihoi (kë)	X	X	O	X
hoihoi	Hoihoi (kë)	X	X	O	X
mao	(Hë) mao	X	X	O	X
mau	(Tü) mau (hou)	X	X	O	X
mau	(Kë) mau (toi)	X	X	O	X
moutere	Mau	X	O	O	O
moutere	Hoe hë kë	OOO	OOO	O	OOO
moutere	Tae kë	OO	OO	O	XO
tü	Hou/hoe	O/O	O	X	X
tü	Kei	O	O	X	X
tü	Hei	O	O	X	X
tü	Hei	O	O	X	X
matiu	(Tü) mau hoe	XO	OO	X	X
matiu	Tae kë	OO	OO	X	X

Table F - 5Misrecognition results on the pre-recorded Non-Maori Speaker using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	kë kï	O	O	O	X
täne	tae tü (kï)	O	O	O	X
hau	Pao	O	O	O	O
Hou	(Tae) pau	O	O	O	O

	(täne)				
Pao	(Tae) pou	X	O	O	O
Pau	(tae) pou	X	O	O	O
Pou	(Tae) tü	O	O	O	X
pö	(tai) pou	X	O	O	O
Pai	tae (hï)	O	O	O	O
Pae	Tae	O	X	O	X
Kë	(Tae) kë	X	X	O	X
kei	(tae) kï	X	O	O	X
kï	Kï	X	X	X	X
hë	Pae	O	O	O	X
hei	hë (kï)	X	O	O	O
hï	(Tae) kï	O	X	X	X
tae	(tae) hë	X	O	O	X
tai	(Tae) kei	O	O	O	X
mätao	Pae/tae pou	O	O	O	O
mätau	Tae tae/pou	O	XO	O	XO
mätou	tae tae/pou	O	O	O	X
toetoe	tae (kei) tae/pae (kï)	O	O	O	X
toi	tae (pae kï)	X	O	O	X
hoihoi	pae (hë) pai (kï)	O	O	O	X
hoe	(Mau) pae (kë)	O	O	O	O
mao	(hï) pou	O	O	O	O
mau	(Tae) pou	O	O	O	O
moutere	mau kë kë	O	O	O	O
tü	tü	X	X	X	X
matiu	mao (kë) tü	X	O	X	X

Table F - 6: Misrecognition results on the Non-Maori Speaker live when using an average quality microphone in a recording room. HMMs were trained using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	Täne	X	O	O	O
täne	(Tae pao) täne	X	X	O	X
hau	Täne	O	O	O	X
Hou	Tae	O	O	O	X
Pao	Täne	O	O	O	X
Pau	Täne	O	O	O	X
Pou	Tae	O	O	O	X

pö	Täne	O	O	O	X
Pai	Täne	O	O	O	X
Pae	Tënei	O	O	O	O
Kë	Hë	O	X	O	X
kei	Täne	O	O	O	X
kï	Täne	O	O	X	X
hë	Hë	X	X	O	X
hei	Täne	O	O	O	X
hï	Täne	O	O	X	X
tae	(Kë) Täne	X	O	O	O
tai	(Kë) Täne	X	O	O	X
mätao	Kë Tae	O	O	O	X
mätau	(Pae pae) mätao	X	O	O	O
mätou	Mätao (hoe)	X	O	O	O
toetoe	Pae	O	O	O	X
toi	(Hë) Täne	X	O	O	X
hoihoi	Pae pao	O	O	O	O
hoe	Pae (täne)	O	O	O	O
mao	Mätau	X	O	O	O
mau	(Kei) pao	O	O	O	O
moutere	Pae tae pai	O	O	O	X
tü	Kei	O	O	X	X
matiu	Pae tënei	O	O	X	X

Table F - 7: Misrecognition results on the Non-Maori Speaker live when using a broadcasting microphone in a recording room. HMMs were trained using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	täne	X	O	O	X
täne	Tae hë	O	O	O	O
hau	hë	X	O	O	X
Hou	mao	O	O	O	O
Pao	Tae (tü)	O	O	O	X
Pau	Täne	O	O	O	X
Pou	Mätau	O	O	O	X
pö	Tae (tü)	O	O	O	X
Pai	Täne	O	O	O	O
Pae	Täne	O	O	O	O
Kë	tënei	O	O	O	X
kei	Hë/hï	O	O	O	O

kï	Hï (pou)	O	X	X	X
hë	Tënei	O	O	O	O
hei	Hë/hï	X	O	O	X
hï	Hï	X	X	X	X
tae	Täne	X	O	O	O
tai	Täne	X	O	O	O
mätao	Mau täne	X	O	O	O
mätau	Mätao (pae)	X	O	O	O
mätou	Tae pai	O	O	O	X
toetoe	Toi (hë) tae (hei)	X	O	O	O
toi	Pae/pao (hï)	O	O	O	X
hoihoi	Toi (hï) pae (hï)	O	O	O	O
hoe	Pae (hë)	O	O	O	O
mao	Mätau	X	O	O	X
mau	Mätou	X	O	O	X
moutere	Tae tae tënei	O	O	O	X
tü	Tü	X	X	X	X
matiu	Tae tü	O	O	X	X

Table F - 8: Misrecognition results on the Phonetics Expert live when using an average quality microphone in a recording room. HMMs were trained using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	Tü pai	O	O	O	X
täne	täne	X	X	O	X
hau	Tae	O	O	O	X
Hou	Pou	O	X	O	X
Hau	Pae/pou/hou	O	O	O	O
Hou	(Tae) pou	O	X	O	X
Pao	Pau	X	O	O	O
pau	Tae	O	O	O	X
Pou	Pou	X	X	O	X
Pao	Pao	X	X	O	X
Pou	Täne	O	O	O	X
Pö	Toi	O	O	O	X
Pö	(Tae) pao	X	O	O	X
Pai	Tai	O	X	O	X
Pae	kë	O	O	O	X
Pae	kï	O	O	O	X

Kë	tae	O	O	O	X
kï	kï	X	X	X	X
kï	(Tai) kë	X	O	O	X
Kei	Tai	O	O	O	X
hë	Kë	O	X	O	X
Hei	Hei	X	X	O	X
hï	(Tü) kï	O	X	X	X
Tae	Tae (pai)	X	X	O	X
Tai	Tae (kei)	X	O	O	O
Mätao	(Tü täne) mätou	X	O	O	O
Mätau	Tü täne tae kei täne	O	O	O	X
mätou	(Tü) mätau (mau)	X	O	O	O
Toetoe	Tae (tai) toi (hou)	O	O	O	O
Toi	Toi	X	X	O	X
Hoihoi	(Mau) pae (tai) hoe (tai)	O	O	O	O
Hoe	pae (kï)	O	O	O	O
Mao	(Tae) pao	O	X	O	X
Mau	(Tü mätou) hau	O	X	O	X
moutere	(Tü) mao tae kë	O	O	O	O
tü	(Kï) pou	O	O	X	X
matiu	Tae kï hau pou	O	O	X	X

Table F - 9: Misrecognition results on the Phonetics Expert live when using a broadcasting microphone in a recording room. HMMs were trained using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	Tai	O	O	O	O
täne	Hë hou	O	O	O	X
hau	täne	O	O	O	X
Hou	Pao/pou	O	O/X	O	O/X
Pao	Pao	X	X	O	X
Pau	Mao	O	O	O	O
Pou	Mao	O	O	O	O
pö	Mao	O	O	O	O
Pai	(Täne) pai	X	X	O	X

Pae	Täne	O	O	O	O
Kë	tënei	O	O	O	X
kï	Tü	O	O	X	X
kï	Tü	O	O	X	X
hë	Kë	O	X	O	X
hei	tënei	O	X	O	X
hï	Kï	O	O	X	X
tae	täne	X	O	O	O
tai	Tae/pai	X/O	O	O	O
mätao	(Tü) mätao	X	X	O	X
mätau	(Täne) tae kei	O	O	O	X
mätou	(Tü) mätao (pou)	X	O	O	O
toetoe	Toi pae	O	O	O	O
toi	Toi (kei)	X	X	O	X
hoihoi	Hoihoi	X	X	O	X
hoe	Hoe/hei	X	X	O	X
mao	(Tü) mao	X	X	O	X
mau	(Tü) mao (pou)	X	O	O	O
moutere	(Tü) pou kë kei	O	O	O	O
tü	Kei	O	O	X	X
matiu	Tü pae (kï)	O	O	X	X
kei	Hë	O	O	O	O

Table F - 10: Misrecognition results on the Proficient Maori Speaker live when using a broadcasting microphone in a recording room. HMMs were trained using K and R training data

Intended Word	Recognised Word	Misrecognised consonant	Misrecognised Vowel	Confusable Vowel Pair Exists	Misrecognised as the other vowel of the pair
tënei	Tënei	X	X	O	X
täne	Täne	X	X	O	X
hau	Pau	O	X	O	X
Hou	Tai	O	O	O	X
Pao	Mao	O	X	O	X
Pau	Pau	X	X	O	X
Pou	Tae	O	O	O	X
pö	Pao	X	O	O	O
Pai	Pai	X	X	O	O
Pae	Tae	O	X	O	X
Kë	Kei	X	O	O	O
kei	Kei	X	X	O	X
kï	Kï	X	X	X	X
hë	Kë	O	X	O	X
hei	Pae (kï)	O	O	O	X

hï	Kï	O	X	X	X
tae	Pae	O	X	O	X
tai	Pai	O	X	O	X
mätao	Mätao	X	X	O	X
mätau	Mätau (hei)	X	X	O	X
mätou	Mätou	X	X	O	X
toetoe	Toi hoe (kei)	O	O	O	O
toi	Toi (kï)	X	X	O	X
hoihoi	Hoihoi	X	X	O	X
hoe	Pae	O	O	O	X
mao	Mao	X	X	O	X
mau	Pau	O	X	O	X
moutere	Pau kei	O	O	O	X
tü	Tai	X	O	X	X
matiu	Matiu	X	X	X	X

The following tables were made under the assumption that while there could be multiple errors of different types for each word, there could only be one error of each type. For example, the second word, in Table F-11, *täne,* was misrecognised as *Pau hoe*. This misrecognition has two vowel substitution errors and two consonant substitution errors but as there is only one example of this word, it is counted as one vowel substitution error and one consonant substitution error.

Table F - 11: Error Types occurring in each recognized word for the Y group when trained on the K group

WORD	K/Y						
	MISRECOGNISED AS	NUMBER CORRECT	NUMBER OF VOWEL SUBSTITUTIONS	NUMBER OF CONSONANT SUBSTITUTIONS	NUMBER OF INSERTIONS	NUMBER OF DELETIONS	TOTAL
tënei	1x Kë hei 1x (Kei) tënei	0	0	1	1	0	2
täne	1x Pau hoe	0	1	1	0	0	2
hau	1x Pai	0	1	1	0	0	2
Hou	1x Pau 1x Pou	0	1	2	0	0	3
Pao	2x Mao 1x pao (pao)	0	0	2	1	0	3
Pau	1x mau	0	0	1	0	0	1

Pou							
pö							
Pai	2x pae	0	2	0	0	0	2
Pae							
Kë							
kei	1x kï 1x kë 1x tae (kï)	0	3	1	1	0	4
kï	1x kë	0	1	0	0	0	1
hë	4x kë	0	0	4	0	0	4
hei	1x kë 2x hï	0	3	1	0	0	4
hï	4x kï 1x (kë) hï	0	0	4	1	0	5
tae	3x pae	0	0	3	0	0	3
tai	2x pai 1x pai (hoe)	0	0	3	1	0	4
mätao	1x mätou/pao 1x mätao/hou	0	2	2	2	0	6
mätau	1x mätau (tai) 1x mätou (tai)	0	1	0	2	0	3
mätou	1x mätau 1x mätau/pou	0	2	1	1	0	4
toetoe	3x toi toi 1x toi (pao) hoe	0	4	1	1	0	6
toi							
hoihoi							
hoe	1x hoihoi 1x toi 1x hou/toi	0	2	2	2	0	6
mao							
mau							
moutere	1x mau toi 1x mau tënei 1x hou kë kë 1x mau kë	0	4	4	0	2	10
tü	1x kë 2x kï	0	3	3	0	0	6
matiu	1x Toi kï	0	1	1	0	0	2
TOTAL	55	0	31	38	13	2	83

Table F - 12: Error Types occurring in each recognized word for the H group when trained on the R group

WORD	R/H						
	MISRECOGNISED AS	NUMBER CORRECT	NUMBER OF VOWEL SUBSTITUTIONS	NUMBER OF CONSONANT SUBSTITUTIONS	NUMBER OF INSERTIONS	NUMBER OF DELETIONS	TOTAL
tënei	1x Kë tae 1x Kë tae kë 1x Hë toi	0	3	3	1	0	7
täne	1x Hau hë 1x hë 2x täne (kë) 1x tai	0	2	2	2	2	8
hau	1x Tü 1x Pau (kë) 2x Pau	0	1	4	1	0	6
Hou	1x Hoe 1x Tae (kë)	0	2	1	1	0	4
Pao	1x Mao 1x pao/pou 1x mätao (pou)	0	2	1	2	0	5
Pau	4x hau 1x tae/tai	0	1	5	1	0	7
Pou	3x Hou 2x Hau 2x Pau	0	3	5	0	0	8
pö	3x Hou 1x hau 1x pao 1x hë	0	6	5	0	0	11
Pai	1x tae 2x tai 1x pae 1x tae (kei)	0	3	4	1	0	8
Pae	1x tae 1x hau 1x pae (kë) 1x täne	0	1	3	2	0	6
Kë							
kei	1x tai 2x kë 3x hei	0	3	4	0	0	7
kï	2x kë	0	3	1	0	0	4

	1x hou						
hë	4x kë	0	0	4	0	0	4
hei	1x kei 1x tae	0	1	2	1	0	4
hï	2x kei 2x kï 1x kë	0	3	5	0	0	8
tae	1x (pao) pae (kë) 1x pao/pai	2	1	2	1	0	4
tai	1x hoe 1x pai 1x tae	0	2	2	0	0	4
mätao	1x mätou 1x mätau/pou	0	2	1	1	0	4
mätau							
mätou	1x mätau (hë) 1x mätau	0	2	0	1	0	3
toetoe	1x toetoe (pou) 1x toi hoe 1x toi hoe (kë) 3x hoe hoe	0	2	3	2	0	7
toi	1x hoihoi 1x hoe (kei 1x hoe	0	2	3	2	0	7
hoihoi	2x Hoihoi (kë)	0	0	0	2	0	2
hoe							
mao	1x (hë) mao	0	0	0	1	0	1
mau	1x (tü) mau (hou) 1x (kë) mau (toi)	0	0	0	2	0	2
moutere	1x mau 1x hoe kë kë 1x tae kë	0	3	2	0	2	7
tü	1x hou/hoe 1x kei 2x hei	0	4	4	1	0	9
matiu	1x (tü) mau hoe 1x tae kë	0	2	2	1	0	5
TOTAL	96	2	54	68	26	4	152

Table F - 13: Error Types occurring in each recognized word for the Y group when trained on the K and R groups

WORD	KR/Y						
	MISRECOGNISED AS	NUMBER CORRECT	NUMBER OF VOWEL SUBSTITUTIONS	NUMBER OF CONSONANT SUBSTITUTIONS	NUMBER OF INSERTIONS	NUMBER OF DELETIONS	TOTAL
tënei	1x Kë (tü) toi 1x Kë (tai) hoi	0	2	2	2	0	6
täne							
hau	1x hei 1x hou	0	2	0	0	0	2
Hou	1x Pau 1x Pou 1x hau	0	2	2	0	0	4
Pao	2x Mao 1x pao (pao)	1	0	1	1	0	3
Pau	1x mau	0	2	0	0	0	2
Pou	1x pau	0	1	0	0	0	1
pö	1x tü	0	1	1	0	0	2
Pai	1x pae 1x toi 1x hoe	0	3	2	0	0	5
Pae							
Kë							
kei	1x kï 1x tae (kï)	0	2	1	1	0	4
kï	1x kë	0	1	0	0	0	1
hë	2x kë 2x kï	0	2	4	0	0	6
hei	1x kei 1x hoe 2x kï	0	2	3	0	0	5
hï	3x kï 1x kï (kï)	0	0	4	1	0	5
tae	2x pae 1x pai	0	1	3	0	0	4
tai	1x pai 1x hoe	0	1	2	0	0	3
mätao							
mätau	1x mätou (tai) 1x mao pou	0	2	1	1	0	4

	(tai)						
mätou	3x mätau	0	3	0	0	0	3
toetoe	2x toi toi 1x hoe (pao) hoe	0	2	1	1	0	4
toi							
hoihoi	1x hoe hoe	0	1	0	0	0	1
hoe							
mao							
mau	1x pau	0	0	1	0	0	1
moutere	1x hou kë kë	0	1	1	0	0	2
tü	3x kï	0	3	3	0	0	6
matiu	1x kï	0	1	1	0	1	3
TOTAL	51	1	35	33	7	1	76

Table F - 14: Error Types occurring in each recognized word for the H group when trained on the K and R groups

WORD	**KR/H**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
tënei	1x Kë tae 1x Kë toi	0	2	2	0	0	4
täne	2x täne (kë) 1x tae tai 1x hau hë 1x hou	0	3	3	1	1	8
hau	1x hau (tü) 1x tae	0	1	1	1	0	3
Hou	1x Pou 1x Hei 1x hoe	0	2	1	0	0	3
Pao	1x (mätao) pou 1x pao (tae) 1x pao/pou	0	2	0	3	0	5
Pau	1x tae 2x hou 2x hau	0	3	3	0	0	6
Pou	3x hou	0	2	5	0	0	7

	1x tae 1x tü						
pö	1x pö/pou 1x pao 2x hou 1x hau/pao	0	5	3	2	0	10
Pai	2x pae1x hoe 1x pai/pou 1x pai/kei 1x pae/tae 1x pou	0	5	3	3	0	11
Pae							
Kë	1x hë	0	0	1	0	0	1
kei	3x hei	0	0	3	0	0	3
kï	2x kë 1x kei	0	3	0	0	0	3
hë	3x kë 1x (täne) kë	0	0	4	1	0	5
hei							
hï	3x kï 1x kë 1x kei	0	2	5	0	0	7
tae	4x pae 1x pae (kë)	0	0	5	1	0	6
tai	1x tae (pou) 1x pae 2x pai 1x pae (kei)	0	3	4	2	0	9
mätao	1x mätao/pae 3x mätao/pou	0	0	0	4	0	4
mätau	2x mätou	0	2	0	0	0	2
mätou	2x mätao (hoe) 1x mätau 1x mätau (tae)	0	4	0	3	0	7
toetoe	2x hoe hoe 1x toi hoe (pou) 1x toi hoe (kë) 1x toi toi 1x hoe tënei	0	4	4	3	0	11
toi	1x hoihoi 3x hoe (kei)	0	3	4	4	0	11
hoihoi	1x hoe hoe 1x hoihoi (kë) 1x hoihoi (kei)	0	1	0	2	0	3
hoe	1x hoe (kei) 1x hoihoi 1x hoe/toi	0	2	1	3	0	6

mao							
mau	1x (tü) mau (tü) 1x (kë) pae 1x mau (tü)	0	1	1	3	0	5
moutere	1x mau 1x hou/mau tënei 1x mao/hou kë kë 1x tae tënei 1x hoihoi tai kë	0	5	4	3	1	13
tü	1x mau (kë kë) 1x kë 2x kei	0	4	4	1	0	9
matiu	1x (tü) matiu (tae/pae) 1x mätao (kë)	0	1	0	2	0	3
TOTAL	101	0	60	61	42	2	165

Table F - 15: Error Types occurring in each recognized word for the pre-recorded non-Maori speaker when trained on the K and R groups

WORD	**KR/Non-Maori Speaker Pre-Recorded**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
tënei	1x kë kï	0	1	1	0	0	2
täne	1x tae tü (kï)	0	1	1	1	0	3
hau	1x pao	0	1	1	0	0	2
Hou	1x (tae pau (täne)	0	1	1	1	0	3
Pao	1x (tae) pou	0	1	0	1	0	2
Pau	1x (tae) pou	0	1	0	1	0	2
Pou	1x (tae) tü	0	1	1	1	0	3
pö	1x (tae) pou	0	1	0	1	0	2
Pai	1x tae (hï)	0	1	1	1	0	3
Pae	1x tae	0	0	1	0	0	1
Kë	1x (tae) kë	0	0	0	1	0	1
kei	1x (tae) kï	0	1	0	1	0	2
kï		1	0	0	0	0	0

hë	1x pae	0	1	1	0	0	2
hei	1x hë (kï)	0	1	0	1	0	2
hï	1x (tae) kï	0	0	1	1	0	2
tae	1x tae (hë)	0	0	0	1	0	1
tai	1x tae (kei)	0	1	0	1	0	2
mätao	1x pae/tae pou	0	1	1	1	0	3
mätau	1x tae tae/pou	0	1	1	1	0	3
mätou	1x tae tae/pou	0	1	1	1	0	3
toetoe	1x tae (kei) 1x Tae/pae (kï)	0	1	1	1	0	3
toi	1x tae (pae kï)	0	1	0	1	0	2
hoihoi	1x pae (hë) pai (kï)	0	1	1	1	0	3
hoe	1x (mau) pae (kë)	0	1	1	1	0	3
mao	1x (hï) pou	0	1	1	1	0	3
mau	1x (tae) pou	0	1	1	1	0	3
moutere	1x mau kë kë	0	1	1	0	0	2
tü		1	0	0	0	0	0
matiu	1x mao (kë) tü	0	1	0	1	0	2
TOTAL	29	2	24	18	23	0	65

Table F - 16: Error Types occurring in each recognized word for the live non-Maori speaker with an average microphone when trained on the K and R groups

WORD	**KR/Non-Maori Speaker Live with Average Microphone**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
tënei	1x täne	0	1	0	0	0	1
täne	1x (tae pao) täne	0	0	0	1	0	1
hau	1x täne	0	1	1	1	0	3
Hou	1x tae	0	1	1	0	0	2
Pao	1x täne	0	1	1	1	0	3
Pau	1x täne	0	1	1	1	0	3
Pou	1x tae	0	1	1	0	0	2
pö	1x täne	0	1	1	1	0	3

Pai	1x täne	0	1	1	1	0	3
Pae	1x tënei	0	1	1	1	0	3
Kë	1x hë	0	0	1	0	0	1
kei	1x täne	0	1	1	1	0	3
kï	1x täne	0	1	1	1	0	3
hë		1	0	0	0	0	0
hei	1x täne	0	1	1	1	0	3
hï	1x täne	0	1	1	1	0	3
tae	1x (Kë) täne	0	0	1	1	0	2
tai	1x (Kë) täne	0	0	1	1	0	2
mätao	1x Kë tae	0	1	1	0	0	3
mätau	1x (Pae pae) mätao	0	1	0	1	0	2
mätou	1x mätao (hë)	0	1	0	1	0	2
toetoe	1x pae	0	1	1	0	1	2
toi	1x (hë) täne	0	1	1	1	0	3
hoihoi	1x pae pao	0	1	1	0	0	2
hoe	1x pae (täne)	0	1	1	1	0	3
mao	1x mätau	0	1	1	1	0	3
mau	1x (kei) pao	0	1	1	1	0	3
moutere	1x pae tae pai	0	1	1	1	0	3
tü	1x kei	0	1	1	0	0	2
matiu	1x pai tënei	0	1	1	1	0	3
TOTAL	29	1	25	24	9	1	59

Table F - 17: Error Types occurring in each recognized word for the live non-Maori speaker with a broadcasting microphone in a recording room when trained on the K and R groups

WORD	**KR/Non-Maori Speaker Live with Broadcasting Microphone in Recording Room**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
tënei	1x täne	0	1	0	0	0	1
täne	1x tae hë	0	1	1	0	0	2
hau	1x hë	0	1	0	0	0	1

Hou	1x mao	0	1	1	0	0	2
Pao	1x tae (tü)	0	1	1	1	0	3
Pau	1x täne	0	1	1	1	0	3
Pou	1x mätau	0	1	1	1	0	3
pö	1x tae (tü)	0	1	1	1	0	3
Pai	1x täne	0	1	1	1	0	3
Pae	1x täne	0	1	1	1	0	3
Kë	1x tënei	0	1	1	1	0	3
kei	1x hë/hï	0	1	1	1	0	3
kï	1x hï (pou)	0	0	1	1	0	2
hë	1x tënei	0	1	1	1	0	3
hei	1x hë/hï	0	1	1	1	0	3
hï		1	0	0	0	0	0
tae	1x täne	0	0	1	1	0	2
tai	1x täne	0	0	1	1	0	2
mätao	1x mau täne	0	1	0	1	0	2
mätau	1x mätao (pae)	0	1	0	1	0	2
mätou	1x tae pai	0	1	1	0	0	2
toetoe	1x toi (hë) tae (hei)	0	1	0	1	0	2
toi	1x pae/pao (kï)	0	1	1	1	0	3
hoihoi	1x toi (hi) pae (hi)	0	1	1	1	0	3
hoe	1x pae (hë)	0	1	1	1	0	3
mao	1x mätau	0	1	0	1	0	2
mau	1x mätou	0	1	0	1	0	2
moutere	1x tae tae tënei	0	1	1	1	0	2
tü		1	0	0	0	0	0
matiu	1x tae tü	0	1	1	0	0	2
TOTAL	28	2	25	21	22	0	68

Table F - 18: Error Types occurring in each recognized word for the live phonetics expert with an average microphone in a recording room when trained on the K and R groups

WORD	KR/Phonetics Expert Live with Average Microphone in Recording Room						
	MISRECOGNISED AS	NUMBER CORRECT	NUMBER OF VOWEL SUBSTITUTIONS	NUMBER OF CONSONANT SUBSTITUTIONS	NUMBER OF INSERTIONS	NUMBER OF DELETIONS	TOTAL
tënei	1x tü pai	0	1	1	0	0	2
täne		1	0	0	0	0	0
hau	1x tae 1x pae/pou/ hou	0	1	1	1	0	3
Hou	1x pou 1x (tae) pou	0	0	2	1	0	3
Pao	1x pau	1	1	0	0	0	1
Pau	1x tae	0	1	1	0	0	2
Pou	1x täne	1	1	1	1	0	3
pö	1x toi 1x (tae) pao	0	2	2	1	0	5
Pai	1x tai	0	0	1	0	0	1
Pae	1x kë 1x kï	0	2	2	0	0	4
Kë	1x tae	0	1	1	0	0	2
kei	1x tai	0	1	1	0	0	2
kï	1x (tai) kë	1	1	0	1	0	2
hë	1x kë	0	0	1	0	0	1
hei		1	0	0	0	0	0
hï	1x (tü) kï	0	0	1	1	0	2
tae	1x tae (pai)	0	0	0	1	0	1
tai	1x tae (kei)	0	1	0	1	0	2
mätao	1x (tü täne) mätou	0	1	0	1	0	2
mätau	1x tü täne tae kei täne	0	1	1	1	0	3
mätou	1x (tü) mätau (mau)	0	1	1	1	0	3
toetoe	1x tae (tai) toi (hou)	0	1	0	1	0	2
toi		1	0	0	0	0	0
hoihoi	1x (mai) pae (tai) hoe (tai)	0	1	1	1	0	3

hoe	1x pae (kï)	0	1	1	1	0	3
mao	1x (tae) pao	0	0	1	1	0	2
mau	1x (tü mätou) hau	0	1	1	1	0	3
moutere	1x (tü) mao tae kë	0	1	1	1	0	3
tü	1x (kï) pou	0	1	1	1	0	3
matiu	1x Tae kï hau pou	0	1	1	1	0	3
TOTAL	31	6	23	24	19	0	66

Table F - 19: Error Types occurring in each recognized word for the live phonetics expert with a broadcasting microphone in a recording room when trained on the K and R groups

WORD	**KR/Phonetics Expert Live with Broadcasting Microphone in Recording Room**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
tënei	1x tai	0	1	0	0	1	2
täne	1x hë hou	0	1	1	0	0	2
hau	1x täne	0	1	1	1	0	3
Hou	1x pao/pou	0	1	1	1	0	3
Pao		1	0	0	0	0	0
Pau	1x mao	0	1	1	0	0	2
Pou	1x mao	0	1	1	0	0	2
pö	1x mao	0	1	1	0	0	2
Pai	1x (täne) pai	0	0	0	1	0	1
Pae	1x täne	0	0	1	1	0	2
Kë	1x tënei	0	1	1	1	0	3
kei	1x tü 1x hë	0	2	2	0	0	4
kï	1x tü	0	1	1	0	0	2
hë	1x kë	0	0	1	0	0	1
hei	1x tënei	0	0	1	1	0	23
hï	1x kï	0	0	1	0	0	1

tae	1x täne	0	0	1	1	0	2
tai	1x tae/pai	0	1	1	1	0	3
mätao	1x (tü) mätao	0	0	0	1	0	1
mätau	1x (täne) tae kei	0	1	1	1	0	3
mätou	1x (tü) mätao (pou)	0	1	1	1	0	3
toetoe	1x toi pae	0	1	1	1	0	3
toi	1x toi (kei)	0	0	0	1	0	1
hoihoi		1	0	0	0	0	0
hoe	1x hoe/hei	0	1	0	1	0	2
mao	1x (tü) mao	0	0	0	1	0	1
Mau	1x (tü) mao (pou)	0	1	1	1	0	3
Moutere	1x pou kë kei	0	1	1	1	0	3
Tü	1x kei	0	1	1	0	0	2
Matiu	1x tü pae (kï)	0	1	1	1	0	3
TOTAL	29	2	20	23	18	1	62

Table F - 20: Error Types occurring in each recognized word for the live proficient Maori Speaker with a broadcasting microphone in a recording room when trained on the K and R groups

WORD	**KR/Proficient Maori Speaker Live with Broadcasting Microphone in Recording Room**						
	MISRECOGNISED AS	**NUMBER CORRECT**	**NUMBER OF VOWEL SUBSTITUTIONS**	**NUMBER OF CONSONANT SUBSTITUTIONS**	**NUMBER OF INSERTIONS**	**NUMBER OF DELETIONS**	**TOTAL**
Tënei		1	0	0	0	0	0
Täne		1	0	0	0	0	0
Hau	1x pau	0	0	1	0	0	1
Hou	1x tai	0	1	1	0	0	2
Pao	1x mao	0	0	1	0	0	1
Pau		1	0	0	0	0	0
Pou	1x tae	0	1	1	0	0	2
Pö	1x pao	0	1	0	0	0	1
Pai		1	0	0	0	0	0
Pae	1x tae	0	0	1	0	0	1

Kë	1x kei	0	1	0	0	0	1
Kei		1	0	0	0	0	0
Kï		1	0	0	0	0	0
Hë	1x kë	0	0	1	0	0	1
Hei	1x pae (kï)	0	1	1	1	0	3
Hï	1x kï	0	0	1	0	0	1
Tae	1x pae	0	0	1	0	0	1
Tai	1x pai	0	0	1	0	0	1
Mätao		1	0	0	0	0	0
Mätau	1x mätau (hei)	0	0	0	1	0	1
Mätou		1	0	0	0	0	0
Toetoe	1x toi hoe (kei)	0	1	1	1	0	3
Toi	1x toi (kï)	0	0	0	1	0	1
Hoihoi		1	0	0	0	0	0
Hoe	1x pae	0	1	1	0	0	2
Mao		1	0	0	0	0	0
Mau	1x pau	0	0	1	0	0	1
Moutere	1x pau kei	0	1	1	0	1	3
Tü	1x tai	0	1	0	0	0	1
Matiu		1	0	0	0	0	0
TOTAL		11	8	14	4	1	27