

---

---

# Speech recognition based on classification

word pairs versus a general model

---

---

Author

Tim Haverkamp  
*Artificial Intelligence Radboud University,  
Nijmegen*

Supervised by

Makiko Sadakata,  
Jana Krutwig  
*Artificial Intelligence Radboud University,  
Nijmegen*

JUNE 2016

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Classification</b>	<b>4</b>
<b>3</b>	<b>Manual Annotating</b>	<b>4</b>
3.1	Sound as an image . . . . .	5
3.2	Recognizing vowels and consonants . . . . .	6
3.3	Processing annotation . . . . .	6
3.3.1	Extracting vowels . . . . .	6
3.3.2	Formants . . . . .	6
<b>4</b>	<b>Hidden Markov Model Toolkit (HTK)</b>	<b>8</b>
<b>5</b>	<b>Training and test data</b>	<b>10</b>
<b>6</b>	<b>Results</b>	<b>11</b>
6.1	Word pairs & General model . . . . .	11
6.2	Binomial test . . . . .	12
<b>7</b>	<b>Conclusion &amp; Discussion</b>	<b>13</b>
7.1	Discussion . . . . .	13
7.2	Conclusion . . . . .	14
<b>8</b>	<b>References</b>	<b>15</b>
<b>A</b>	<b>Log probabilities word-pair model</b>	<b>16</b>
<b>B</b>	<b>Log probabilities general model</b>	<b>17</b>

## List of Figures

1	All the words spoken by Female 5 with first and second formant track . . . . .	5
2	All the words spoken by Male 3 with first and second formant band	5
3	Results of annotating the first four words from Female five by three different people . . . . .	7
4	The word 'fan' by Male 3 . . . . .	7
5	Annotated image of the word 'fan' by Male 3 . . . . .	8
6	Processing stages of HTK [6] . . . . .	9
7	Monophone, Biphone and Triphone HMM's for the word bat.[2]	9
8	The visual representation of the word pan by Male speaker 3 . .	11
9	The visual representation of the word pen by Male speaker 3 . .	11
10	Manual analysis of Male 2 . . . . .	14

## List of Tables

1	Results for the word-pair model on each of the word-pairs . . . .	12
2	Results for the general model on the word-pairs . . . . .	13

## Abstract

The classification of a word-pair specific model versus a general model is being evaluated. Word-pairs are difficult for non-native speakers and automatic speech recognition to recognize and classify as they are very similar in features. The classification is carried out by a toolkit that uses hidden markov models on the audio files. Manual analysis is shown to indicate how difficult speech is to comprehend and how it differs per person. Significant differences on performance have been found between the general model and the word-pair model carried out by the toolkit.

## 1 Introduction

A lot of work has been done regarding speech recognition and its uses are present in our every day lives. We use it on our phones, computer and devices keep coming that use a form of speech recognition. Although we use it with devices, we all know that sometimes it just does not work the way we want to. This is because speech is personal and unique, everybody has a different way of speaking. One might speak differently among other people, is tired or has a cold which will result in different speech. This can make the job harder for speech recognition software as it sometimes requires perfect pronunciation of words. A few methods have thus been suggested to let speech recognition perform better [7].

Also people can have problems recognizing words that are similar in form and speech. Take the words pan and pen for example, these three letter words look quite similar and also sound alike depending on the speaker. This might pose as a problem for a non-native speaker. A person might have a higher or lower fundamental frequency (The speed with which the vocal cords vibrate[5]), caused by the amount of open space between the vocal cords and the tension that is being used on the vocal cords. Besides sounding the same, the properties of the sounds also look a lot like each other, when looking at figure 8 and figure 9 quite some similarities can be seen. Not only does it look a lot like each other but the properties are also quite similar. Taking a look at the formant tracks (harmonic of a note augmented by a resonance [12]) which are used to distinguish vowels, you see that they lie closely together which makes it harder for a system and even a human to recognize a vowel as the correct one if the speech of the vowels are overlapping. That is why training a model specifically for the classification of vowels is so important. It can make the difference between words and meaning of a sentence. In section 3 an explanation will be given of how the sounds look alike and what problems that might give.

In this thesis an attempt will be made to identify whether the Hidden Markov Model is suited for the classification of /a/ and /e/ vowel words and if word pair specific models need to be trained or if a general model is sufficient to classify speech. In short the answers will be given for the following questions:

- *Is the Hidden Markov Model toolkit sufficient for vowel classification?*
- *Is a word pair specific model needed to recognize the words correctly or is a general model sufficient?*

By answering these questions I hope to show that although speech recognition is far, still a lot of progress needs to be made before being able to recognize speech fully. Section 2 gives a short indication of how the classification is achieved, section 3 describes the problems when manually annotating speech. Section 4 discusses the toolkit and gives a general idea of how the system is used. The data is discussed in Section 5 and results of the tests are shown in Section 6. Finally Section 7 discusses the results and a conclusion for the thesis will be made.

## 2 Classification

The workings of the toolkit that was used to classify the data is described in section 4. a short description of how the classification is achieved will be discussed in this section.

After running the data through the hidden markov models the output of the system is checked against labeled files of the test data. Not only does it check to see if it is different than what was labeled but it will also check for differences within the word. So instead of looking at sentences and misplaced words it looks at the contents of the words. This enables the classification of /a/ and /e/ vowels more clearly. With the help of manually annotated audio files emphasis could be laid upon the difficulty of recognizing speech.

## 3 Manual Annotating

Getting from raw audio files to vowel information requires some steps. One of these steps is annotating the audio files and segmenting them as vowel or consonant. These following chapters will explain the steps that are needed to accomplish this and also what to do with the extracted information. In a previous section it was explained that the vowels /a/ and /e/ look quite similar but there are also differences within each spoken word. Figure 2 and figure 1 show these similarities within themselves and between each other. The first and second formant of Female 5 is very spread out and the /a/ and /e/ vowel are mixed together whereas with Male 3 you have more clustered formant tracks and the first and second formant are almost fully separated.

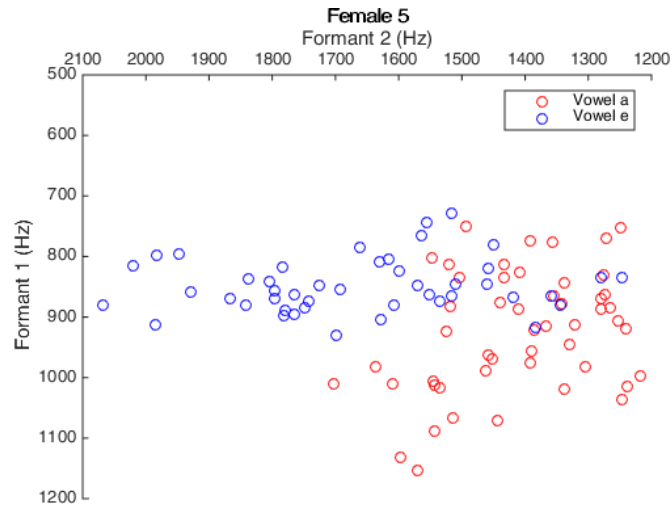


Figure 1: All the words spoken by Female 5 with first and second formant track

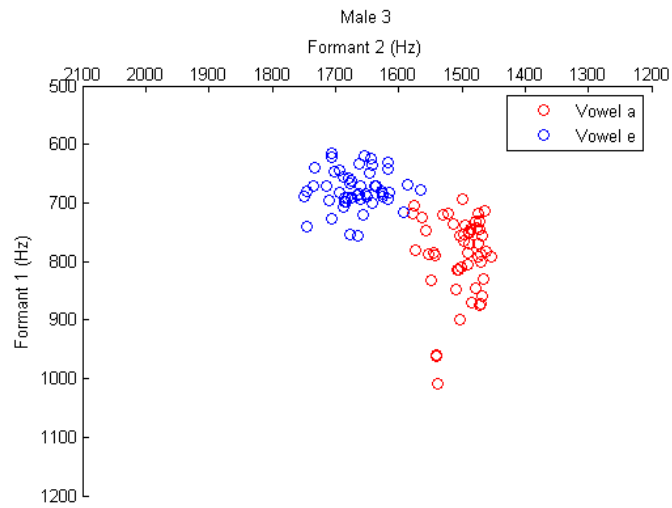


Figure 2: All the words spoken by Male 3 with first and second formant band

### 3.1 Sound as an image

Manually segmenting is done in a few steps. Software called Praat is used for this purpose, it can manipulate and analyze speech and show visually how an

audio signal looks. The raw audio files are displayed when first opening the sound files in Praat as seen in figure 4. The top image is an oscillogram and the bottom image is a spectrogram. The blue lines in the spectrogram represents the fundamental frequency of the person and the red dots are the formant tracks.

## 3.2 Recognizing vowels and consonants

In this example we have the word fan with f being a fricative (a consonant with an obstruction on its way out). The transition from a fricative to a vowel is clearly visible in both images of figure 4. In the oscillogram normal noise can be seen and then a lot of oscillations. In the spectrogram nothing clearly can be seen and then a lot of black bands(formants) in the image. At this transition a segmentation can be made, the left part being a consonant and the right part a vowel. Now most words have co-articulation meaning that the same phonemes are not identical in every situation. Phonemes are influenced by surrounding phonemes which makes it harder to determine borders in words[13], so listening to the word can be useful to better determine the borders. When done with the 'f' and 'a' letters, the 'n' needs to be segmented. However the image does not show a clear border as it was with the 'f' and 'a' but more of a decreasing signal. Again listening to the word and making a segmentation approximately at the end of the black bands would get a somewhat accurate result. The resulting image would then be as figure 5 with 'c' being a consonant and 'v' being a vowel.

## 3.3 Processing annotation

After the annotation is done the segmentation of the vowel is needed to extract all kinds of data from the separate vowel sound file.

### 3.3.1 Extracting vowels

Before extracting all kinds of information from the vowel in a word it needs to be isolated from the consonants. This is done by a script that searches for the letter 'v' in the text grids that was made by annotating the word as seen in figure 5. This is then stored in a separate sound file where information can be extracted from the vowel.

### 3.3.2 Formants

Formants are a way of distinguishing vowels from each other and most of the time the first and second formants are used to distinguish them. However sometimes the third formant can give some extra information. Results of the first, second and third formant are being discussed further on. In a previous part co-articulation was mentioned, co-articulation has an influence on the information of a phoneme. Trying to eliminate information from certain consonants in a vowel is done by a cut in a vowel to make sure that the formant tracks are from the vowel itself. This is done by requesting the starting time and end

time of a vowel and determining the first and third quarter. Extracting the three formant tracks is done by cutting out the middle part between the first and third quarter. This formant information is then stored in a separate text file that can be processed later on. As seen in figure 3 Manual Annotating can be quite inconsistent as speech is subjective. The smallest mistake can lead to a different understanding of a word which in turn can lead to a difference in understanding a sentence. Therefore toolkits exist to automatically recognize speech, one of these is called Hidden Markov Model Toolkit. The next section discusses this toolkit in more detail.

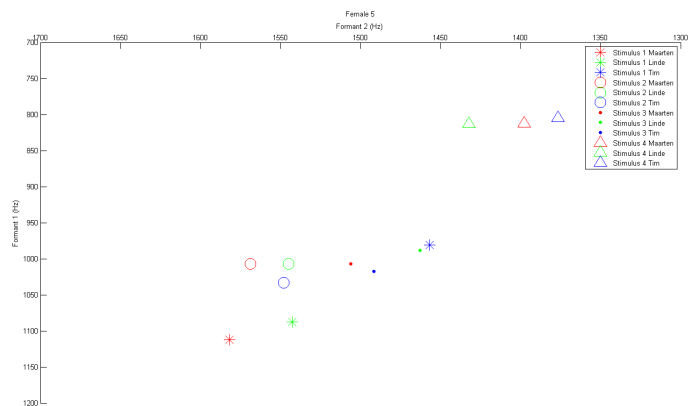


Figure 3: Results of annotating the first four words from Female five by three different people

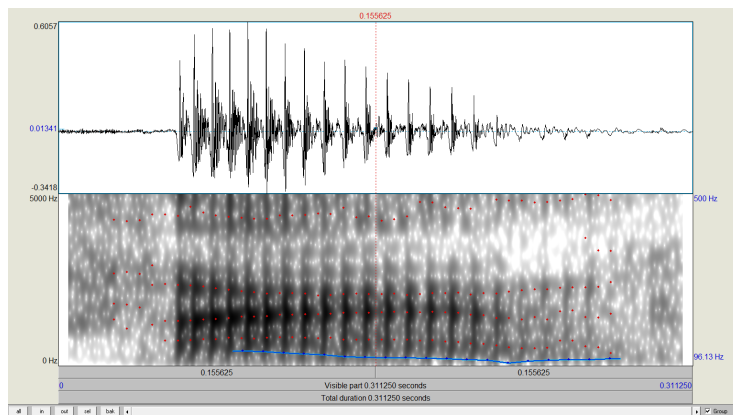


Figure 4: The word 'fan' by Male 3



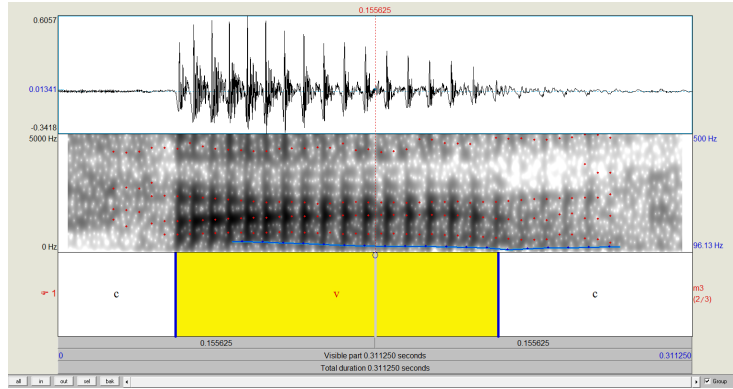


Figure 5: Annotated image of the word 'fan' by Male 3

## 4 Hidden Markov Model Toolkit (HTK)

The Hidden Markov Model Toolkit (HTK) is designed for speech recognition, although it can be used for different purposes.[14] It does this by using/training Hidden Markov Models (HMM's) on the training data and then classifying on the remaining test data. The reason why HTK is used is for its low complexity, great Real-Time performance and the usage of a lot of parameters.[1] The hmm's are initially represented by states and a transition matrix per monophone. These are determined by looking at the global means and variances in the audio files. These hmm's keeps changing when re-estimating. [4]

Using the HTK requires some steps to be made, these steps are all documented in a manual made by the creators of the toolkit. The steps for classifying words with the help of HTK are as follows. A grammar of the target words needs to be made and be transformed into a word network. A dictionary of a list of words needs to be made and a list of monophones that are used in the dictionary are listed in a separate file. For HTK to train the Markov Models it needs to know what the audio files mean, so transcriptions of those audio files need to be put into a file. Then the audio training files need to be transformed into the right format in order for the Hidden Markov models to be able to use it. A prototype model is needed to indicate how the monophones(one individual phone, a single character of a word represented by a symbol. [5]) are represented, vector size and a matrix are indicated. After this is done a command needs to be run that updates the model into the first Hidden Markov model and a re-estimation is done two more times. Silence models are added into the Hidden Markov definition model to absorb noise that is present in the audio files. The system needs to update the Markov models two more times. Now the system needs to realign its training data, the most common pronunciation of the grammar words are now determined and the system is re-estimated two more times. Triphones are made from the monophones that were determined earlier on. This is done

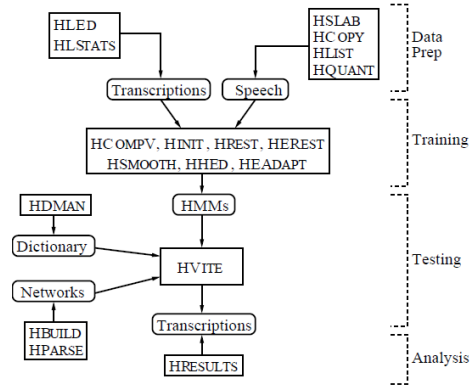


Figure 6: Processing stages of HTK [6]

to account for context in which a monophone can exist. A monophone can sound differently depending on surrounding monophones, thus a triphone list is made to make sure the system takes it into account.[5] The system needs to be re-estimated two more times but now with the triphone list. The triphones now need to be tied together to share data this will make sure that if one triphone in one instance is changed then the other tied triphones are also changed. A figure explaining the monophone and triphone model can be found at figure 7. A decision tree is created to deal with triphones that have not been seen before and a re-estimation is done twice again. Now the Hidden Markov Models are run on the test data and the output of the HMM's are checked with transcription files made for the test data. A more detailed step by step is found in the manual of the toolkit[6] and the general processing steps of HTK can be seen in figure 6.

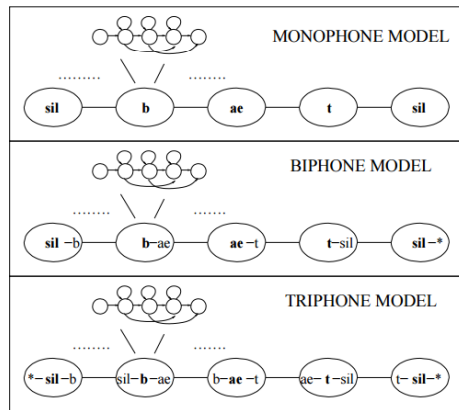


Figure 7: Monophone, Biphone and Triphone HMM's for the word bat.[2]

## 5 Training and test data

The data consist of audio files from nine different British speakers. These nine people (four female and five male) all spoke ten words multiple times. There are two word pairs within these ten words, the word pairs that are used in the data are fan-fen, jam-gem, ham-hem, man-men and pan-pen. The only difference are the vowels /a/ and /e/. The problematic part about these kinds of words are the similarities between vowels and the pronunciation variation of those words. A few characteristics when listening to speech can shed some light on the similarities. One of these characteristics is formant tracks. Formants are a way of distinguishing vowels from each other and are visually represented as bands in a spectrogram. Figure 8 and 9 show these similarities. Although differences can be spotted in the spectrogram (bottom part) the form is the same. Some black bands grouped together and a space in between the top and bottom part. The differences between pan and pen regarding the oscillogram (top part) are also small, the form, again, is the same. These differences between words are thus hard to distinguish sometimes. Besides differences between words there are also differences within word pairs. Every word is pronounced differently which makes it harder for Automatic Speech Recognition (ASR) to learn words/speech.[3] When speaking the person might get tired or there is something in their throat or there is an accent in play which is only used in certain circumstances. These differences might lead to formant tracks of the a vowel looking more like formant tracks of the e vowel and vice versa. Perhaps for native speakers it won't be a problem to comprehend what is said but for non-native and also for speech recognition this might pose as a problem. The data will thus be tested on their likeliness of it being a word with an a or an e vowel.

The division of the data set had been set to 144 training and 36 test data. This was done because of the available data for the word-pair model. Every word was spoken about ten times and with nine subjects the total equals 90. Word-pairs were tested so 90 times 2 equals 180 audio files for a word-pair. To equalize this for the general model a randomization had been done. For the nine subjects four words were chosen randomly and of those chosen words the first, last and middle three had been picked out. This had been done because of how the word-pair model was tested. Results of these tests will be discussed in the next section.

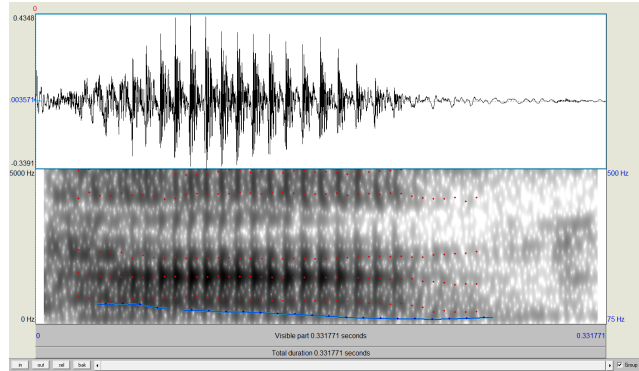


Figure 8: The visual representation of the word pan by Male speaker 3

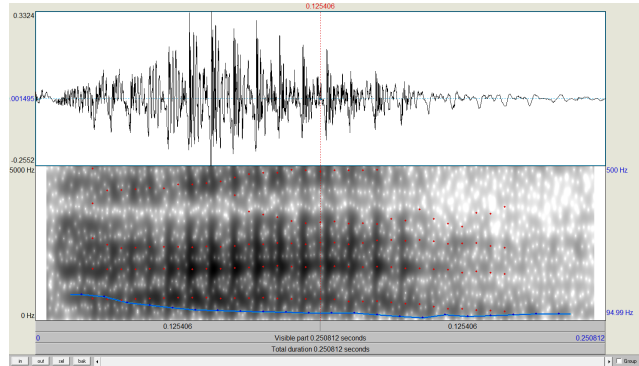


Figure 9: The visual representation of the word pen by Male speaker 3

## 6 Results

This section will show the results obtained by running the toolkit several times. Each individual word-pair was trained and a general model was trained on which the word-pairs were tested. This results in 30 trials in total.

### 6.1 Word pairs & General model

This section will show the results for testing the word pairs with HTK. The correctness percentage is calculated by a formula as seen in equation 1 .[6] Insertions are additions to certain words that it will not expect it to be there, deletions are vowels or consonants that are missing and substitutions are replacements that are made in words. Substitutions can be seen as a combination of a deletion and an insertion.[5] The high number of insertions is caused by the addition of word-start and word-end that is defined in the grammar file. The

Fan-Fen	Run 1	Run 2	Run 3	Average
Correctness	97.22	100	94.44	97.22
Hits	35	36	34	35
Deletions	0	0	0	0
Substitutions	1	0	2	1
Insertions	72	72	72	72
Jam-Gem	Run 1	Run 2	Run 3	Average
Correctness	94.44	97.22	97.22	96.29
Hits	34	35	35	34.67
Deletions	0	0	0	0
Substitutions	2	1	1	1.33
Insertions	72	72	72	72
Pan-Pen	Run 1	Run 2	Run 3	Average
Correctness	91.67	94.44	91.67	92.59
Hits	33	34	33	33.33
Deletions	0	0	0	0
Substitutions	3	2	3	2.67
Insertions	72	72	72	72

Ham-Hem	Run 1	Run 2	Run 3	Average
Correctness	94.22	94.44	94.44	94.38
Hits	34	34	34	34
Deletions	0	0	0	0
Substitutions	2	2	2	2
Insertions	72	72	72	72
Man-Men	Run 1	Run 2	Run 3	Average
Correctness	86.11	97.22	88.89	90.74
Hits	31	35	32	32.67
Deletions	0	0	0	0
Substitutions	5	1	4	3.33
Insertions	72	72	72	72

Table 1: Results for the word-pair model on each of the word-pairs

results of the HTK can be found in tables 1 and 2. Overall the word-pair model outperforms the general model by several percentages, in some cases even more than 40 percent.

$$Correctness = \frac{Number of files - deletions - substitutions}{Number of files} * 100 \quad (1)$$

## 6.2 Binomial test

For a statistical measure on the results a binomial test was used to compare the two models. The formula for the binomial test can be seen in equation 2.  $n$  stands for the number of times the models produces different output and  $s$  means that model A performs better than model B. Looking at the results the number of different output between the two models is fifteen,  $n$ , and  $s$  for the success rate is also fifteen. Deciding that model A and model B have the same chance of performance,  $p$  and  $q$ , the equation can be filled in. The result is a significance level of  $p \approx 0.00003$  so  $p < 0.005$ . This might indicate a significance of the word-pair model over the general model. Section 7.1 discusses the implications of the significance level that was found.

$$\frac{n!}{s!(n-s)!} p^s * q^{n-s} \quad (2)$$

Fan-Fen	Run 1	Run 2	Run 3	Average
Correctness	77.78	75	80.56	77.78
Hits	28	27	29	28
Deletions	0	0	0	0
Substitutions	8	9	7	8
Insertions	72	72	72	72
Jam-Gem	Run 1	Run 2	Run 3	Average
Correctness	66.67	72.22	69.44	69.44
Hits	24	26	25	25
Deletions	0	0	0	0
Substitutions	12	10	11	11
Insertions	72	72	72	72
Pan-Pen	Run 1	Run 2	Run 3	Average
Correctness	75	75	66.67	72.33
Hits	27	27	24	26
Deletions	0	0	0	0
Substitutions	9	9	12	10
Insertions	72	72	72	72

Ham-Hem	Run 1	Run 2	Run 3	Average
Correctness	52.78	58.33	52.78	54.63
Hits	19	21	19	19.67
Deletions	0	0	0	0
Substitutions	17	15	17	16.33
Insertions	72	72	72	72
Man-Men	Run 1	Run 2	Run 3	Average
Correctness	66.67	63.89	66.67	65.74
Hits	24	23	24	23.67
Deletions	0	0	0	0
Substitutions	12	13	12	12.33
Insertions	72	72	72	72

Table 2: Results for the general model on the word-pairs

## 7 Conclusion & Discussion

The results from the previous sections and the procedure that was used will be discussed in the following section. A conclusion and future work will be formulated in the last section.

### 7.1 Discussion

The randomization of the data as mentioned in section 5 could have caused an unfair balance of represented words in the training model which could have resulted in a smaller classification percentage as seen in the general model. It should be noted that from the output given by the toolkit a lot of substitutions were present in the male subjects. When looking at the manual analysis of male two, as seen in figure 10, it shows that there is a lot of overlap between the /a/ and /e/ vowel with regards to the formants. When looking back at figure 2, it has a more clear distinction between the /a/ and /e/ vowel than Male 2. After each change on the hmm's, re-estimation was done twice in order to fit the new data. The handbook that comes with the toolkit talks about running it only twice because overfitting might occur. Notice should be taken on the amount of runs that was done in trying to randomize the data to determine if the outcomes were not coincidental. More runs per word-pair could be done to further test if the results were coincidental or not. For better results in the models a lot more data could have been added, more subjects resulting in a more diverse spread of speech between subjects or more audio files per spoken word resulting in a bigger pool to draw the data from. These recommendations could provide for

some future work and possibly provide some insight into the improvement in the general model.

Appendices A and B contain the log likelihoods from the first runs. It is obtained by running a re-estimation from one hidden markov model to another. Between certain transitions of markov models extra steps are taken such as triphone inclusion to account for the co-articulation. This can cause a slight increase in the log probability but overall it decreases over the amount of runs. This is an indication that the system is still improving because you want to minimize the negative log probability.

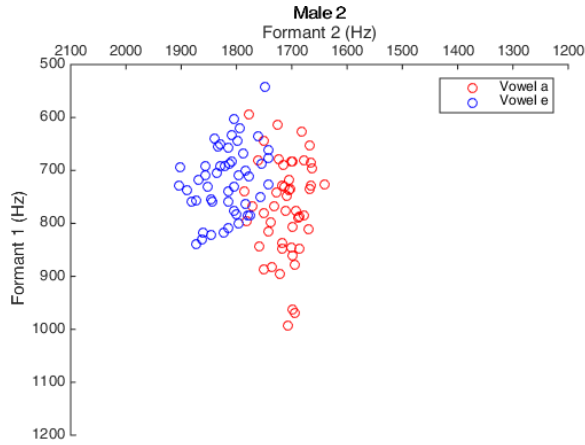


Figure 10: Manual analysis of Male 2

## 7.2 Conclusion

Tests have been done to show whether a word-pair model outperforms a general model and by how much. The results showed that performance of the general model is less compared to the word-pair model, as was expected. When looking closer at the output of the toolkit it showed that words like 'men' were being classified as 'hem' and the switch from the vowel /a/ to the vowel /e/ occurred several times especially in the male subjects. The output of the word-pairs showed that although correctness percentages were high, still a lot of substitutions were present. The HTK seems to be sufficient in vowel classification in a general model with the set-up that was used in this thesis. Even though the general model is sufficient in the vowel classification a word-pair model outperforms the general model by quite some percentages. Co-articulation could be a significant factor in the misclassification of the vowels. Further research and development of the HTK is needed in order to eliminate or decrease the influence of co-articulation.

## 8 References

### References

- [1] Picone, J *Fundamentals of speech recognition: A short course*, Institute for Signal and Information Processing Mississippi State University, 1996
- [2] Resch, B *Automatic Speech Recognition with HTK*, Signal Processing and Speech Communication Laboratory
- [3] Strik, H & Cucchiaroni, C *Modeling pronunciation variation for ASR: A survey of the literature*, Elsevier Speech communication 29, 1999
- [4] Young, S.J *The HTK hidden Markov model toolkit: Design and philosophy*, Cambridge University Engineering Department, 1994
- [5] Rietveld, A.C.M & Heuven van, V.J *Algemene fonetiek*, publisher Coutinho, 2009
- [6] Young, S.J & Evermann, G & Gales, M & Hain, T &, Kershaw, D & Liu, X & Moore, G & Odell, J & Ollason, D & Povey, D & Valtchev, V & Woodland, P *The HTK book*, Cambridge University, December 1995
- [7] Maheswari, N.U & Kabilan, A.P & Venkatesh, R *Speaker independent speech recognition system based on phoneme identification,” Computing, Communication and Networking*, ICCCN 2008. International Conference on, St. Thomas, 2008
- [8] Campbell, J & Tremain, T *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86*, 1986
- [9] Jeans, J.H. *Science & Music*, reprinted by Dover, 1968
- [10] Catford, J.C. *A Practical Introduction to Phonetics*, Oxford University Press p 161, 1988
- [11] Hardcastle, William J & Hewlit, Nigel *Coarticulation theory, data and techniques*, Cambridge University Press p 7, version 2006
- [12] Jeans, J.H. *Science & Music*, reprinted by Dover, 1968
- [13] Hardcastle, William J & Hewlit, Nigel *Coarticulation theory, data and techniques*, Cambridge University Press p 7, version 2006
- [14] Hidden Markov Model Toolkit: <http://htk.eng.cam.ac.uk/>



## A Log probabilities word-pair model

Fan-Fen	log-prob	Ham-Hem	log-prob	Jam-Gem	log-prob	Man-Men	log-prob
0 $\rightarrow$ 1	-76.28240	0 $\rightarrow$ 1	-73.33686	0 $\rightarrow$ 1	-81.18511	0 $\rightarrow$ 1	-72.82254
1 $\rightarrow$ 2	-69.74753	1 $\rightarrow$ 2	-67.96547	1 $\rightarrow$ 2	-70.39219	1 $\rightarrow$ 2	-67.25169
2 $\rightarrow$ 3	-65.67259	2 $\rightarrow$ 3	-64.74414	2 $\rightarrow$ 3	-66.60828	2 $\rightarrow$ 3	-63.77585
5 $\rightarrow$ 6	-64.98499	5 $\rightarrow$ 6	-64.14442	5 $\rightarrow$ 6	-65.94303	5 $\rightarrow$ 6	-63.33562
6 $\rightarrow$ 7	-64.70282	6 $\rightarrow$ 7	-63.91624	6 $\rightarrow$ 7	-5.60908	6 $\rightarrow$ 7	-63.15378
7 $\rightarrow$ 8	-64.53597	7 $\rightarrow$ 8	-63.89396	7 $\rightarrow$ 8	-65.77761	7 $\rightarrow$ 8	-63.13424
8 $\rightarrow$ 9	-64.44211	8 $\rightarrow$ 9	-63.81770	8 $\rightarrow$ 9	-65.62128	8 $\rightarrow$ 9	-63.07560
10 $\rightarrow$ 11	-64.39193	10 $\rightarrow$ 11	-63.77825	10 $\rightarrow$ 11	-65.44793	10 $\rightarrow$ 11	-63.05858
11 $\rightarrow$ 12	-64.06853	11 $\rightarrow$ 12	-63.35394	11 $\rightarrow$ 12	-65.06462	11 $\rightarrow$ 12	62.71914
13 $\rightarrow$ 14	-64.15338	13 $\rightarrow$ 14	-63.41823	13 $\rightarrow$ 14	-64.99039	13 $\rightarrow$ 14	-62.77974
14 $\rightarrow$ 15	-64.14133	14 $\rightarrow$ 15	-63.39981	14 $\rightarrow$ 15	-64.73772	14 $\rightarrow$ 15	-62.76940
Pan-Pen	log-prob						
0 $\rightarrow$ 1	-74.83671						
1 $\rightarrow$ 2	-69.00440						
2 $\rightarrow$ 3	-65.79454						
5 $\rightarrow$ 6	-65.32206						
6 $\rightarrow$ 7	-65.05559						
7 $\rightarrow$ 8	-65.01386						
8 $\rightarrow$ 9	-64.92105						
10 $\rightarrow$ 11	-64.88251						
11 $\rightarrow$ 12	-64.52097						
13 $\rightarrow$ 14	-64.52029						
14 $\rightarrow$ 15	-64.49288						

## B Log probabilities general model

Fan-Fen	log-prob	Ham-Hem	log-prob	Jam-Gem	log-prob	Man-Men	log-prob
0 $\rightarrow$ 1	-75,99839	0 $\rightarrow$ 1	-76,01473	0 $\rightarrow$ 1	-76,00263	0 $\rightarrow$ 1	-76,06895
1 $\rightarrow$ 2	-70,34889	1 $\rightarrow$ 2	-70,36452	1 $\rightarrow$ 2	-70,35005	1 $\rightarrow$ 2	-70,42509
2 $\rightarrow$ 3	-66,48389	2 $\rightarrow$ 3	-66,47247	2 $\rightarrow$ 3	-66,51213	2 $\rightarrow$ 3	-66,58592
5 $\rightarrow$ 6	-65,86084	5 $\rightarrow$ 6	-65,83824	5 $\rightarrow$ 6	-65,89141	5 $\rightarrow$ 6	-65,95325
6 $\rightarrow$ 7	-65,64851	6 $\rightarrow$ 7	-65,62870	6 $\rightarrow$ 7	-65,68909	6 $\rightarrow$ 7	-65,74350
7 $\rightarrow$ 8	-65,59434	7 $\rightarrow$ 8	-65,86751	7 $\rightarrow$ 8	-65,64052	7 $\rightarrow$ 8	-65,68960
8 $\rightarrow$ 9	-65,52961	8 $\rightarrow$ 9	-65,63111	8 $\rightarrow$ 9	-65,57834	8 $\rightarrow$ 9	-65,62673
10 $\rightarrow$ 11	-65,49768	10 $\rightarrow$ 11	-65,51365	10 $\rightarrow$ 11	-65,54522	10 $\rightarrow$ 11	-65,59434
11 $\rightarrow$ 12	-63,48711	11 $\rightarrow$ 12	-63,43818	11 $\rightarrow$ 12	-63,54017	11 $\rightarrow$ 12	-63,63046
13 $\rightarrow$ 14	-65,23597	13 $\rightarrow$ 14	-65,20264	13 $\rightarrow$ 14	-65,28768	13 $\rightarrow$ 14	-65,33201
14 $\rightarrow$ 15	-65,17055	14 $\rightarrow$ 15	-65,13469	14 $\rightarrow$ 15	-65,22108	14 $\rightarrow$ 15	-65,26494
Pan-Pen	log-prob						
0 $\rightarrow$ 1	-76,12565						
1 $\rightarrow$ 2	-70,46331						
2 $\rightarrow$ 3	-66,07800						
5 $\rightarrow$ 6	-65,99362						
6 $\rightarrow$ 7	-65,79935						
7 $\rightarrow$ 8	-65,75842						
8 $\rightarrow$ 9	-65,69875						
10 $\rightarrow$ 11	-65,65978						
11 $\rightarrow$ 12	-63,67181						
13 $\rightarrow$ 14	-65,34755						
14 $\rightarrow$ 15	-65,27649						