

PHONEME BASED SPEECH TO TEXT TRANSLATION  
SYSTEM FOR MALAYSIAN ENGLISH  
PRONUNCIATION

SATHEES KUMAR NATARAJ

UNIVERSITI MALAYSIA PERLIS  
2012

© This item is protected by original copyright



**Phoneme Based Speech to Text Translation System for  
Malaysian English Pronunciation**

by

**SATHEES KUMAR NATARAJ**  
**(0830610271)**

A thesis submitted in fulfillment of the requirements for the degree of  
Master of Science (Mechatronic Engineering)

**School of Mechatronic Engineering**  
**UNIVERSITI MALAYSIA PERLIS**

2012

## ACKNOWLEDGEMENT

I would like to extend my sincere gratitude to the Vice Chancellor of UniMAP, **Brigedier Jeneral Dato' Prof. Dr. Kamarudin b. Hussin** for his constant encouragement throughout my study.

I extend my wholehearted thankfulness to **Prof. Dr. Sazali Bin Yaacob**, Head of Intelligent Signal Processing Research Cluster, UniMAP for his unconditional support and motivation right through my research work and who has also co-supervised my research work.

I would like to express my profound gratefulness to **Prof. Madya. Dr. Abdul Hamid Adom**, Dean, School of Mechatronic Engineering, UniMAP for providing support and encouragement throughout my research work.

I am very fortunate to have a kindhearted person to guide me throughout my research and would like to show my heartfelt sincere gratitude to my beloved mentor, and first Supervisor **Prof. Madya. Dr. Paulraj Murugesu Pandiyan**, Mechatronic Engineering Programme, School of Mechatronic Engineering, UniMAP for his invaluable guidance, support, and enthusiasm. I am greatly indebted for his inspiration and thirst for knowledge and research which helped me to groom and nurture myself throughout my research work. His continuous motivation helped me to complete my research successfully.

I also wish to extend my sincere gratitude to **Prof. Madya. Ahmad Nazri Bin Abdullah**, Assoc. Prof. Sultan Abdul Hamid College, Alor Setar, Kedah, Malaysia, for his funding support and who has also co-supervised my research work.

I would like to thank, the members of staff, School of Mechatronic Engineering, Research and Development, Library, ICT, Bendahari and Postgraduate studies for their kind and needful assistance, encouragement and support during my research work.

I would like to express my deep gratitude to the Government of Malaysia for providing me an opportunity to accomplish my research and studies.

I have been fortunate to have many friends who cherish me despite my eccentricities. I thank Mr. Srinivasan S, Mr. Dinesh Kumar, Mr. Siva Sundhar, Mr. Jagadeesh, Mr. Syahrull Hi-fi Ahmad Jamil and Mr. Erdy Sulino Mohd Muslim Tan for continuous and excellent guidance and advice during my years of M.Sc. study. I also thank all the members of Intelligent Signal Processing Research Cluster and all the members of the Research Laboratory and other research clusters for their endless support and motivation.

Finally, I would like to thank my parents for their unconditional support, both financially and emotionally throughout my research work. In particular, the patience and understanding shown by my mother, father and sister during the honors year is greatly appreciated. I know, at times, my temper is particularly trying.

## TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE NO</b>
<b>THESIS DECLARATION</b>	i
<b>ACKNOWLEDGEMENT</b>	ii
<b>TABLE OF CONTENTS</b>	iv
<b>LIST OF TABLES</b>	ix
<b>LIST OF FIGURES</b>	xvii
<b>LIST OF ABBREVIATIONS</b>	xxiii
<b>LIST OF SYMBOLS</b>	xxv
<b>ABSTRAK</b>	xxvii
<b>ABSTRACT</b>	xxviii
<b>CHAPTER 1 INTRODUCTION</b>	
1.1 Preamble	1
1.2 Motivation towards the research	1
1.3 Problem statement	2
1.4 Research approach	3
1.5 Research objective and significance	3
1.6 Research methodology	4
1.7 Thesis organization	6
<b>CHAPTER 2 LITERATURE REVIEW ON PHONEME BASED SPEECH TO TEXT TRANSLATION SYSTEM</b>	
2.1 Introduction	8
2.2 Malaysian english pronunciation	8
2.3 Speech signal	9
2.3.1 Voiced/Unvoiced classification of speech signal	9

2.4	Speech recognition systems	12
2.4.1	Current state of isolated word recognition systems	14
2.5	Phonemes	17
2.5.1	History of phonemes	18
2.5.2	Approach towards phoneme recognition	20
2.6	Motivation towards phoneme based speech to text translation	23
2.7	Summary	25
<b>CHAPTER 3 ISOLATED WORD SPEECH DATA COLLECTION AND PRE-PROCESSING</b>		
3.1	Introduction	26
3.2	Selection of non native malaysian speakers	26
3.3	Preparation of wordlist	28
3.4	Experimental setup	32
3.5	Data collection procedure	33
3.6	Malaysian English speech database	35
3.6.1	Vowel Class Word Database (VCWD)	35
3.6.2	Phoneme class word database (PCWD)	40
3.7	Pre-Processing of the isolated word speech signals	41
3.8	Summary	43
<b>CHAPTER 4 FEATURE EXTRACTION AND DATA PROCESSING</b>		
4.1	Introduction	44
4.2	Fuzzy voiced/unvoiced classification	44
4.2.1	Frame blocking	47
4.2.2	Energy and change in energy features	50
4.2.3	Fuzzzy voiced/unvoiced classifier	54

4.3	Fuzzy voice classifier computation results	61
4.4	Phoneme segmentation	71
4.5	Feature extraction methods for isolated word and phoneme classification	75
4.5.1	Mel Frequency Cepstral Co-Efficient (MFCC)	77
4.5.2	Linear Predictive Co-efficient (LPC)	81
4.5.3	Perceptual Linear Prediction (PLP)	86
4.5.4	Relative Spectral Transform - Perceptual Linear Prediction (RASTA-PLP)	90
4.6	Data preprocessing	93
4.6.1	Data normalization and randomization	93
4.6.2	Output labeling	95
4.7	Summary	99
<b>CHAPTER 5 SPEECH TO TEXT TRANSLATION SYSTEM USING MULTILAYER NEURAL NETWORKS</b>		
5.1	Introduction	101
5.2	Synopsis of artificial neural network	101
5.3	Motivations for the choice of MLNN	102
5.4	Neural network training and testing	103
5.5	Architecture and design of multilayer feed forward neural network	103
5.5.1	MLNN architecture for VCWD database	104
5.5.2	MLNN architecture for PCWD database	105
5.5.3	Phoneme based speech to text translation system	108
5.6	Training multilayer feed forward neural network	110
5.6.1	Levenberg marquardt training algorithm	110
5.6.2	Choosing neural network parameters	111

5.7	Isolated word classification results using VCWD	114
5.7.1	MFCC classification results	115
5.7.2	Comparison of MFCC architecture	120
5.7.3	LPCC classification results	121
5.7.4	Comparison of LPC architecture	125
5.7.5	RASTA-PLPC classification results.	126
5.7.6	Comparison of RASTA-PLPC architecture	130
5.7.7	PLPC classification results	131
5.7.7	Comparison of PLPC architecture	135
5.8	Comparison of mean classification accuracy	136
5.9	Isolated word classification results using PCWD	137
5.9.1	MLNN fusion architecture classification results	137
5.9.2	Comparison of LPC architecture	142
5.9.3	Comparison of mean classification accuracy	143
5.10	Phoneme classification results using PCWD	144
5.10.1	MLNN fusion architecture classification results	144
5.10.2	Comparison of LPC architecture	149
5.10.3	Comparison of mean classification accuracy	150
5.11	MATLAB graphical user interface for the speech to text translation	151
5.12	Summary	153
<b>CHAPTER 6 RESULTS DISCUSSION AND CONCLUSION</b>		
6.1	Introduction	154
6.2	Discussion	154
6.3	Thesis conclusion	159
6.4	Future work	160

<b>REFERENCES</b>	161
<b>APPENDIX A – ISOLATED WORD SPEECH SIGNAL PLOT USING PCWD DATABASE</b>	173
<b>APPENDIX B – CONFUSION MATRIX</b>	202
<b>LIST OF PUBLICATIONS</b>	255

© This item is protected by original copyright

**LIST OF TABLES**

<b>NO.</b>		<b>PAGE</b>
3.1	Vowel classifications	29
3.2	Isolated wordlist based on phonemic variation	30
3.3	English word list with phonemic variations	31
4.1	Rules formulated for the fuzzy voice classifier	57
4.2	Fuzzy voice classifier results for the VCWD database	61
4.3	Fuzzy voice classifier results for PCWD (Vowels class)	62
4.4	Fuzzy voice classifier results for PCWD (Diphthong class)	63
4.5	Fuzzy voice classifier results for PCWD (Consonant class)	64
4.6	Fuzzy voice classifier results for PCWD (Affricatives class)	65
4.7	Fuzzy voice classifier results for PCWD (Fricatives class)	66
4.8	Fuzzy voice classifier results for PCWD (Semivowels and glides class)	67
4.9	Fuzzy voice classifier results for PCWD (Nasals class)	68
4.10	Fuzzy voice classifier results for PCWD (Consonant class)	69
4.11	Phonemes distribution in the recorded isolated words speech signals	71
4.12	Phonemes classes and its time duration	73
4.13	Phonemes classes and its time duration	73
4.14	Combination of activation function in hidden and output neurons	94
4.15	Data normalization combinations	95
4.16	Target output mapping (logistic sigmoidal normalization)	96
4.17	Target output mapping (hyperbolic tangent sigmoidal normalization)	96
4.18	Target output mapping for affricatives using (logistic sigmoidal normalization)	97
4.19	Target output mapping for diphthongs using (logistic sigmoidal normalization)	97

4.20	Target output mapping for fricatives using (logistic sigmoidal normalization)	97
4.21	Target output mapping for liquids using (logistic sigmoidal normalization)	98
4.22	Target output mapping for nasals using (logistic sigmoidal normalization)	98
4.23	Target output mapping for semivowels and glides using (logistic sigmoidal normalization)	98
4.24	Target output mapping for stops using (logistic sigmoidal normalization)	98
4.25	Target output mapping for vowels using (logistic sigmoidal normalization)	99
5.1	MLNN classification performance for MFCC features (60 % Samples)	117
5.2	MLNN classification performance for MFCC features (70 % Samples)	118
5.3	MLNN classification performance for MFCC features (80 % Samples)	119
5.4	Network training parameters and training time (MFCC features)	120
5.5	MLNN classification performance for LPCC features (60 % Samples)	122
5.6	MLNN classification performance for LPCC features (70 % Samples)	123
5.7	MLNN classification performance for LPCC features (80 % Samples)	124
5.8	Network training parameters and training time (LPC features)	125
5.9	MLNN classification performance for RASTA-PLP features (60 % Samples)	127
5.10	MLNN classification performance for RASTA-PLP features (70 % Samples)	128
5.11	MLNN classification performance for RASTA-PLP features (80 % Samples)	129
5.12	Network training parameters and training time (RASTA-PLPC features)	130
5.13	MLNN classification performance for PLPC features (60 % Samples)	132
5.14	MLNN classification performance for PLPC features (70 % Samples)	133
5.15	MLNN classification performance for PLPC features (80 % Samples)	134

5.16	Network training parameters and training time (PLPC features)	135
5.17	Mean classification accuracy for the MLNN models using VCWD database	136
5.18	MLNN classification performance for LPCC features (60 % Samples)	139
5.19	MLNN classification performance for LPCC features (70 % Samples)	140
5.20	MLNN classification performance for LPCC features (80 % Samples)	141
5.21	Network training parameters and training time (LPCC features)	142
5.22	Mean classification accuracy for the MLNN models	143
5.23	MLNN classification performance for LPCC features (60 % Samples)	146
5.24	MLNN classification performance for LPCC features (70 % Samples)	147
5.25	MLNN classification performance for LPCC features (80 % Samples)	148
5.26	Network training parameters and training time (LPCC features)	149
5.27	Mean classification accuracy for the MLNN models	151
B.1	Confusion matrix for the 60 % of MFCC feature using VCWD	202
B.2	Confusion matrix for the 60 % of MFCC feature using VCWD	202
B.3	Confusion matrix for the 60 % of MFCC feature using VCWD	203
B.4	Confusion matrix for the 60 % of MFCC feature using VCWD	203
B.5	Confusion matrix for the 60 % of MFCC feature using VCWD	204
B.6	Confusion matrix for the 70 % of MFCC feature using VCWD	204
B.7	Confusion matrix for the 70 % of MFCC feature using VCWD	205
B.8	Confusion matrix for the 70 % of MFCC feature using VCWD	205
B.9	Confusion matrix for the 70 % of MFCC feature using VCWD	206
B.10	Confusion matrix for the 70 % of MFCC feature using VCWD	206
B.11	Confusion matrix for the 80 % of MFCC feature using VCWD	207
B.12	Confusion matrix for the 80 % of MFCC feature using VCWD	207

B.13	Confusion matrix for the 80 % of MFCC feature using VCWD	208
B.14	Confusion matrix for the 80 % of MFCC feature using VCWD	208
B.15	Confusion matrix for the 80 % of MFCC feature using VCWD	209
B.16	Confusion matrix for the 60 % of LPCC feature using VCWD	209
B.17	Confusion matrix for the 60 % of LPCC feature using VCWD	210
B.18	Confusion matrix for the 60 % of LPCC feature using VCWD	210
B.19	Confusion matrix for the 60 % of LPCC feature using VCWD	211
B.20	Confusion matrix for the 60 % of LPCC feature using VCWD	211
B.21	Confusion matrix for the 70 % of LPCC Feature using VCWD	212
B.22	Confusion matrix for the 70 % of LPCC feature using VCWD	212
B.23	Confusion matrix for the 70 % of LPCC feature using VCWD	213
B.24	Confusion matrix for the 70 % of LPCC feature using VCWD	213
B.25	Confusion matrix for the 70 % of LPCC feature using VCWD	214
B.26	Confusion matrix for the 80 % of LPCC feature using VCWD	214
B.27	Confusion matrix for the 80 % of LPCC feature using VCWD	215
B.28	Confusion matrix for the 80 % of LPCC feature using VCWD	215
B.29	Confusion matrix for the 80 % of LPCC feature using VCWD	216
B.30	Confusion matrix for the 80 % of LPCC feature using VCWD	216
B.31	Confusion matrix for the 60 % of RASTA-PLPC feature using VCWD	217
B.32	Confusion matrix for the 60 % of RASTA-PLPC feature using VCWD	217
B.33	Confusion matrix for the 60 % of RASTA-PLPC feature using VCWD	218
B.34	Confusion matrix for the 60 % of RASTA-PLPC Feature using VCWD	218
B.35	Confusion matrix for the 60 % of RASTA-PLPC feature using VCWD	219
B.36	Confusion matrix for the 70 % of RASTA-PLPC feature using VCWD	219
B.37	Confusion matrix for the 70 % of RASTA-PLPC feature using VCWD	220

B.38	Confusion matrix for the 70 % of RASTA-PLPC feature using VCWD	220
B.39	Confusion matrix for the 70 % of RASTA-PLPC feature using VCWD	221
B.40	Confusion matrix for the 70 % of RASTA-PLPC feature using VCWD	221
B.41	Confusion matrix for the 80 % of RASTA-PLPC feature using VCWD	222
B.42	Confusion matrix for the 80 % of RASTA-PLPC feature using VCWD	222
B.43	Confusion matrix for the 80 % of RASTA-PLPC feature using VCWD	223
B.44	Confusion matrix for the 80 % of RASTA-PLPC feature using VCWD	223
B.45	Confusion matrix for the 80 % of RASTA-PLPC feature using VCWD	224
B.46	Confusion matrix for the 60 % of PLPC feature using VCWD	224
B.47	Confusion matrix for the 60 % of PLPC feature using VCWD	225
B.48	Confusion matrix for the 60 % of PLPC feature using VCWD	225
B.49	Confusion matrix for the 60 % of PLPC feature using VCWD	226
B.50	Confusion matrix for the 60 % of PLPC feature using VCWD	226
B.51	Confusion matrix for the 70 % of PLPC feature using VCWD	227
B.52	Confusion matrix for the 70 % of PLPC feature using VCWD	227
B.53	Confusion matrix for the 70 % of PLPC feature using VCWD	228
B.54	Confusion matrix for the 70 % of PLPC feature using VCWD	228
B.55	Confusion matrix for the 70 % of PLPC feature using VCWD	229
B.56	Confusion matrix for the 80 % of PLPC feature using VCWD	229
B.57	Confusion matrix for the 80 % of PLPC feature using VCWD	230
B.58	Confusion matrix for the 80 % of PLPC feature using VCWD	230
B.59	Confusion matrix for the 80 % of PLPC feature using VCWD	231
B.60	Confusion matrix for the 80 % of PLPC feature using VCWD	231
B.61	Confusion matrix for the classification of isolated words based on vowel classes	232

B.62	Confusion matrix for the classification of isolated words based on vowels classes	232
B.63	Confusion matrix for the classification of isolated words based on vowels classes	233
B.64	Confusion matrix for the classification of isolated words based on diphthongs classes	233
B.65	Confusion matrix for the classification of isolated words based on diphthongs classes	234
B.66	Confusion matrix for the classification of isolated words based on diphthongs classes	234
B.67	Confusion matrix for the classification of isolated words based on consonants classes	235
B.68	Confusion matrix for the classification of isolated words based on consonants classes	235
B.69	Confusion matrix for the classification of isolated words based on consonants classes	236
B.70	Confusion matrix for the classification of isolated words based on consonants classes	236
B.71	Confusion matrix for the classification of isolated words based on consonants classes	237
B.72	Confusion matrix for the classification of isolated words based on consonants classes	237
B.73	Confusion matrix for the classification of isolated words based on affricatives classes	238
B.74	Confusion matrix for the classification of isolated words based on affricatives classes	238
B.75	Confusion matrix for the classification of isolated words based on affricatives classes	239
B.76	Confusion matrix for the classification of isolated words based on fricatives classes	239
B.77	Confusion matrix for the classification of isolated words based on fricatives classes	240
B.78	Confusion matrix for the classification of isolated words based on fricatives classes	240

B.79	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	241
B.80	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	241
B.81	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	242
B.82	Confusion matrix for the classification of isolated words based on nasals classes	242
B.83	Confusion matrix for the classification of isolated words based on nasals classes	243
B.84	Confusion matrix for the classification of isolated words based on nasals classes	243
B.85	Confusion matrix for the classification of phonemes based on affricatives classes	244
B.86	Confusion matrix for the classification of isolated words based on affricatives classes	244
B.87	Confusion matrix for the classification of isolated words based on affricatives classes	244
B.88	Confusion matrix for the classification of isolated words based on diphthongs classes	245
B.89	Confusion matrix for the classification of isolated words based on diphthongs classes	245
B.90	Confusion matrix for the classification of isolated words based on diphthongs classes	246
B.91	Confusion matrix for the classification of isolated words based on fricatives classes	246
B.92	Confusion matrix for the classification of isolated words based on fricatives classes	247
B.93	Confusion matrix for the classification of isolated words based on fricatives classes	247
B.94	Confusion matrix for the classification of isolated words based on liquids classes	248

B.95	Confusion matrix for the classification of isolated words based on liquids classes	248
B.96	Confusion matrix for the classification of isolated words based on liquids classes	249
B.97	Confusion matrix for the classification of isolated words based on nasals classes	249
B.98	Confusion matrix for the classification of isolated words based on nasals classes	250
B.99	Confusion matrix for the classification of isolated words based on nasals classes	250
B.100	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	251
B.101	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	251
B.102	Confusion matrix for the classification of isolated words based on semi vowels and glides classes	251
B.103	Confusion matrix for the classification of isolated words based on stops classes	252
B.104	Confusion matrix for the classification of isolated words based on stops classes	252
B.105	Confusion matrix for the classification of isolated words based on stops classes	253
B.106	Confusion matrix for the classification of isolated words based on vowels classes	253
B.107	Confusion matrix for the classification of isolated words based on vowels classes	254
B.108	Confusion matrix for the classification of isolated words based on vowels classes	254

## LIST OF FIGURES

NO.		PAGE
1.1	Representation of the proposed research methodology	5
2.1	Speech recognition methods	13
2.2	Classification of phonemes	18
3.1	Distribution of malaysian speakers for the isolated word speech data collection	28
3.2	Graphical User Interface (GUI) for the isolated word speech data acquisition	34
3.3	The cardinal vowel chart	36
3.4	Waveform and spectrum representation of the word “beet” and “bit”	37
3.5	Waveform and spectrum representation of the word “boot” and “book”	37
3.6	Waveform and spectrum representation of the word “bait” and “bet”	38
3.7	Waveform and spectrum representation of the word “boat” and “bought”	38
3.8	Waveform and spectrum representation of the word “but”, “bat” and “pot”	39
3.9	Recorded isolated word speech data for the word “pot”	40
3.10	Typical isolated word speech signal of the word ‘yard’	41
3.11	Typical normal isolated word speech signal and the corresponding pre-emphasized signal	43
4.1	A typical recorded continuous speech signal	48
4.2	An isolated word in the recorded speech signal before fuzzy voice segmentation	48
4.3	An isolated word speech signal blocked with frame size of 256 samples and an overlap of 50 percentage	49
4.4	An isolated word speech signal blocked with frame size of 512 samples and an overlap of 50 percentage	50
4.5	Isolated word speech signal recorded for 15 seconds at 16 kHz.	51
4.6	Frame energy, change in energy distribution of the recorded speech signal	51

4.7	Frame energy mapping on the voiced portion of	52
4.8	Flowchart for frame blocking, frame energy and change in energy computation process	53
4.9	Block diagram of the proposed fuzzy voice classifier	54
4.10	Membership function for frame energy	56
4.11	Membership function for change in energy	56
4.12	Membership function for energy in the second frame	56
4.13	Membership function for fuzzy voiced/ unvoiced output	58
4.14	Fuzzy output for the recorded speech signal	58
4.15	A typical extracted isolated voiced portion from the continuous isolated words.	59
4.16	Flowchart for fuzzy voiced/unvoiced classifier	60
4.17	A typical representation of the isolated word “Bear”	74
4.18	A typical representation of phoneme “b”	74
4.19	A typical representation of the phoneme “eh”	75
4.20	A typical representation of phoneme “r”	75
4.21	The isolated word speech signal before and after pre-emphasis	76
4.22	Feature extraction methods	77
4.23	Block diagram of the MFCC processor	78
4.24	The mel-frequency cepstral coefficients	79
4.25	Flowchart for the MFCC feature extraction procedure	81
4.26	Block diagram of the LPC feature extraction	82
4.27	The linear predictive coefficients	83
4.28	Flowchart for the LPC feature extraction procedure	85
4.29	Block diagram of a PLP processor	86
4.30	Spectrogram of the perceptual linear predictive coefficients	87

4.31	Flowchart for the PLP feature extraction procedure	89
4.32	Spectrogram of the perceptual linear predictive coefficients	90
4.33	Flowchart for the RASTA-PLP feature extraction procedure	92
5.1	MLNN architecture for VCWD database	105
5.2	MLNN architecture for PCWD database (Isolated word Classification)	106
5.3	MLNN architecture for PCWD database (Phoneme Classification)	108
5.4	Architecture of the speech to text translation system	109
5.5	Logistic sigmoid activation function	112
5.6	Hyperbolic tangent activation function	113
5.7	Systole activation function	114
5.8	A typical representation of graphical user interface for fuzzy voice classifier	152
5.9	A typical representation of speech to text translation system	153
A.1	Waveform and spectrum representation of the word “sit”	173
A.2	Waveform and spectrum representation of the word “may”	173
A.3	Waveform and spectrum representation of the word “bat”	174
A.4	Waveform and spectrum representation of the word “pot”	174
A.5	Waveform and spectrum representation of the word “luck”	174
A.6	Waveform and spectrum representation of the word “good”	175
A.7	Waveform and spectrum representation of the word “ago”	175
A.8	Waveform and spectrum representation of the word “meat”	175
A.9	Waveform and spectrum representation of the word “car”	176
A.10	Waveform and spectrum representation of the word “soft”	176
A.11	Waveform and spectrum representation of the word “girl”	176
A.12	Waveform and spectrum representation of the word “too”	177
A.13	Waveform and spectrum representation of the word “day”	177

A.14	Waveform and spectrum representation of the word “sky”	177
A.15	Waveform and spectrum representation of the word “boy”	178
A.16	Waveform and spectrum representation of the word “beer”	178
A.17	Waveform and spectrum representation of the word “bear”	178
A.18	Waveform and spectrum representation of the word “tour”	179
A.19	Waveform and spectrum representation of the word “go”	179
A.20	Waveform and spectrum representation of the word “cow”	179
A.21	Waveform and spectrum representation of the word “pit”	180
A.22	Waveform and spectrum representation of the word “bit”	180
A.23	Waveform and spectrum representation of the word “time”	180
A.24	Waveform and spectrum representation of the word “door”	181
A.25	Waveform and spectrum representation of the word “cat”	181
A.26	Waveform and spectrum representation of the word “get”	181
A.27	Waveform and spectrum representation of the word “fan”	182
A.28	Waveform and spectrum representation of the word “van”	182
A.29	Waveform and spectrum representation of the word “think”	182
A.30	Waveform and spectrum representation of the word “that”	183
A.31	Waveform and spectrum representation of the word “send”	183
A.32	Waveform and spectrum representation of the word “zip”	183
A.33	Waveform and spectrum representation of the word “man”	184
A.34	Waveform and spectrum representation of the word “nice”	184
A.35	Waveform and spectrum representation of the word “ring”	184
A.36	Waveform and spectrum representation of the word “leg”	185
A.37	Waveform and spectrum representation of the word “rat”	185
A.38	Waveform and spectrum representation of the word “wet”	185

A.39	Waveform and spectrum representation of the word “hat”	186
A.40	Waveform and spectrum representation of the word “yet”	186
A.41	Waveform and spectrum representation of the word “shop”	186
A.42	Waveform and spectrum representation of the word “leisure”	187
A.43	Waveform and spectrum representation of the word “chop”	187
A.44	Waveform and spectrum representation of the word “jump”	187
A.45	Waveform and spectrum representation of the word “joke”	188
A.46	Waveform and spectrum representation of the word “choke”	188
A.47	Waveform and spectrum representation of the word “taint”	188
A.48	Waveform and spectrum representation of the word “take”	189
A.49	Waveform and spectrum representation of the word “tap”	189
A.50	Waveform and spectrum representation of the word “tape”	189
A.51	Waveform and spectrum representation of the word “taste”	190
A.52	Waveform and spectrum representation of the word “cast”	190
A.53	Waveform and spectrum representation of the word “hunt”	190
A.54	Waveform and spectrum representation of the word “coach”	191
A.55	Waveform and spectrum representation of the word “sea”	191
A.56	Waveform and spectrum representation of the word “zone”	191
A.57	Waveform and spectrum representation of the word “thin”	192
A.58	Waveform and spectrum representation of the word “them”	192
A.59	Waveform and spectrum representation of the word “clothe”	192
A.60	Waveform and spectrum representation of the word “shake”	193
A.61	Waveform and spectrum representation of the word “fish”	193
A.62	Waveform and spectrum representation of the word “theme”	193
A.63	Waveform and spectrum representation of the word “both”	194

A.64	Waveform and spectrum representation of the word “bath”	194
A.65	Waveform and spectrum representation of the word “ray”	194
A.66	Waveform and spectrum representation of the word “way”	195
A.67	Waveform and spectrum representation of the word “hay”	195
A.68	Waveform and spectrum representation of the word “once”	195
A.69	Waveform and spectrum representation of the word “one”	196
A.70	Waveform and spectrum representation of the word “wall”	196
A.71	Waveform and spectrum representation of the word “warm”	196
A.72	Waveform and spectrum representation of the word “yank”	197
A.73	Waveform and spectrum representation of the word “yarn”	197
A.74	Waveform and spectrum representation of the word “yard”	197
A.75	Waveform and spectrum representation of the word “moon”	198
A.76	Waveform and spectrum representation of the word “bottom”	198
A.77	Waveform and spectrum representation of the word “sing”	198
A.78	Waveform and spectrum representation of the word “made”	199
A.79	Waveform and spectrum representation of the word “mass”	199
A.80	Waveform and spectrum representation of the word “mind”	199
A.81	Waveform and spectrum representation of the word “bring”	200
A.82	Waveform and spectrum representation of the word “bang”	200
A.83	Waveform and spectrum representation of the word “knee”	200
A.84	Waveform and spectrum representation of the word “knife”	201
A.85	Waveform and spectrum representation of the word “new”	201
A.86	Waveform and spectrum representation of the word “saw”	201

**LIST OF ABBREVIATIONS**

ANN	Artificial Neural Network
ASR	Automatic Speech Recognition
DCT	Discrete Cosine Transform
DTW	Dynamic Time Warping
DWT	Discrete Wavelet Transform
FFT	Fast Fourier Transform
FVD	Fuzzy Voicing Detector
FVD	Fuzzy Voicing Detector
GUI	Graphical User Interface
HATS	Hidden Activation TRAPS
HMM	Hidden Markov Model
LM	Levenberg Marquardt
LPC	Linear Predictive Co-Efficient
LVASR	Large Vocabulary ASR
MFCC	Mel-Frequency Cepstral Coefficients
ML	Maximum Likelihood
MLNN	Multilayer Neural Network
MLP	The Multilayered Perceptron
MMI	Maximal Mutual Information
MSF	Magnitude Sum Function
PCWD	Phonemes Class Word Database
PER	The Phoneme Error Rate
PLP	Perceptual Linear Predictive analysis
RASTA-PLP	Relative Spectra- Perceptual Linear Predictive analysis