International
Journal
Of Advanced Research in Engineering & Management (IJAREM)

Classification of Vowels in Tamil Language Using Vowel Chart

K. Mathan Kumar (Ph.D)
Electronics and Instrumentation Department
B.S.Abdur Rahman University
Chennai, India

Dr. D. Najumnissa Jamal
Professor and Head of EIE Department
B.S.Abdur Rahman University
Chennai, India

Abstract: Tamil is one of the Dravidian languages spoken by Tamil people of India, There are 12 vowels, 18 consonants, 216 consonant vowels and one aaydham in Tamil language, since vowels are generally long in duration compare to consonant sounds and are spectrally well defined, therefore classification of vowels in Tamil language plays a major role in Tamil speech recognition system. One of the simplest ways of Classifying vowel articulatory configuration is in terms of tongue hump position. This can be done by drawing a vowel chart for format frequencies of different Tamil vowel sounds produced by different native talkers.

Index Terms— vowel chart, Tamil vowels, articulatory configuration, format frequencies.

I. INTRODUCTION

Tamil vowels are called uyireluttu (uyir – life, eluttu – letter). The vowels are classified into kuril (short) and netil (long) and two diphthongs vowels. Tamil letters are also called the "Life or Soul letters". Together, with the consonants, they form compound, syllabic letters, that are called us "Living letters". The Tamil vowel sound produced is determined by the position of the tongue, but the position of the jaw, lips, also influence the resulting Tamil vowel sound. The Tamil vowels are generally long in duration as compared to Tamil consonant sounds and are spectrally well defined. As such they are usually easily recognized and therefore contribute significantly to our ability to recognize speech.

Vowel are articulations can be characterized as height of tongue (high, mid, low), front-back of tongue (front, central, back), roundedness of lips (rounded, unrounded) and tenseness. Different vowel qualities are realized in acoustic analysis of vowels by the relating values of the formant frequencies. The vocal tract acts as a resonant cavity and the position of the jaw, lips, and tongue results in producing different format values. The acoustics of vowels can be visualized and studied by using Spectrograms which displays the acoustic energy at frequency and time domain.

The vowel diagram is the diagrammatical representation of format frequency in the F1 – F2 plane. Vowels are distinct from each other based on their spectral properties. Each vowel in the vowel diagram has a unique first and second format frequency; therefore, we can easily classify the vowels by using vowel diagram. Depending on the language used, it can take the form of a triangle or quadrilateral with respect to the vowel position within the diagram.

II. DATA COLLECTION

The speech samples for 10 different native Tamil male and female speakers pronouncing the same Tamil vowels repeatedly for 7 times like that all the Tamil vowels are separately recorded and it is saved under the filename extension WAV, with 44,100 HZ sampling frequency, 16 bits, mono sound by using praat V.5.3.78 software recording tool with the help of a microphone separately attached to the laptop.

III. ACOUSTICS ANALYSIS

The vocal tract acts as a resonant cavity results in producing different vowels with different frequency format values, this can be visualized using spectrograms.

At first Tamil vowel sound was recorded and saved under the filename extension WAV, and then they were translated in to spectrogram for acoustic analysis in the software praat V.5.3.70. Then the format frequencies F1 and F2 of each vowel from the spectrogram are founded, Figure1 shows the spectrogram of Tamil vowel '**ஏ**',

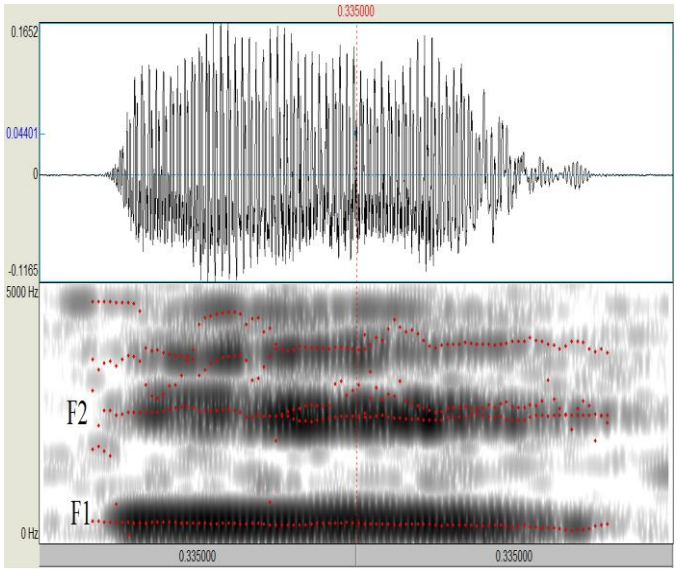

Fig. 1. Spectrogram of Tamil vowel '**ஏ**'

The figure 1 shows how format frequencies is obtained from the single vowel for a male speaker, the darker area in the spectrogram is called as formant represent energy peaks in specific frequency range, F1 is the first dark bar and F2 is the second dark bar, we can obtain the value of F1 and F2 by clicking at the automatically generated red line through the middle of the dark bar, like this format frequencies for all the vowels are found out separately, and then this values are tabulated in table.

IV. TAMIL VOWELS FORMAT FREQUENCY TABULATION

The format frequency obtained by the spectrogram are tabulated by taking average value of F1 and F2 of each male and female speaker pronouncing the same Tamil vowels for 7 times, and then the average value of F1 and F2 for each and every Tamil vowel pronounced by 30 different male and female speaker was calculated and tabulated, and the results are shown in the TABLE I for male and TABLE II for female.

TABLE I. Average format frequency value for 10 male speakers

S.no	Vowels	Average Value of F1	Average Value of F2
1	அ	609.81	874.29
2	ஆ	630.25	905.15
3	இ	333.42	3544.44
4	ஈ	277.33	2615.52
5	உ	331.12	475.03
6	ஊ	369.92	705.95
7	எ	425.07	2307.45
8	ஏ	371.64	2395.85
9	ஐ	425.14	1975.65

10	ஒ	395.49	685.51
11	ஓ	429.71	668.87
12	ஔ	413.89	696.78

TABLE II. Average format frequency value for 10 female speakers

S.no	Vowels	Average Value of F1	Average Value of F2
1	அ	770.86	1123.52
2	ஆ	913.77	1172.7
3	இ	312.98	2642.57
4	ஈ	321.88	2655.82
5	உ	315.4	620.39
6	ஊ	339.65	656.21
7	எ	518.21	2083.989
8	ஏ	545.36	2018.07
9	ஐ	542.42	1823.34
10	ஒ	496.51	947.84
11	ஓ	487.84	917.66
12	ஔ	561.15	853.42

Finally tabulation for the average value of the format frequency of both male and female speaker combined was tabulated by taking average value of the both male and female speaker format frequency value for all the Tamil vowels, and the results are shown in the TABLE III.

TABLE III. Average format frequency value for both male and female speakers combined

S.no	Vowels	Average Value of F1	Average Value of F2
1	அ	690.34	998.91
2	ஆ	772.01	1038.93
3	இ	323.2	3093.51
4	ஈ	299.61	2635.67
5	உ	323.26	547.71
6	ஊ	354.79	681.08
7	எ	471.64	2195.72
8	ஏ	458.5	2206.96
9	ஐ	483.78	1899.5
10	ஒ	446	816.68
11	ஓ	458.78	793.27
12	ஔ	487.52	775.1

From the Tables I and II, we can conclude that the format frequency F1 of the male speaker pronouncing Tamil vowels varies from 277 HZ to 631 HZ and the format frequency F1 of the female speaker pronouncing Tamil vowels varies from 312 HZ to 914 HZ. The format frequency F2 of the male speaker pronouncing Tamil vowels varies from 475 HZ to 3545 HZ and the format frequency F2 of the female speaker pronouncing Tamil vowels varies from 620 HZ to 2643 HZ.

V. TAMIL VOWEL DIAGRAM

The vowel diagram is the diagrammatical representation of format frequency in the F1 – F2 plane. Depending upon the position of a Tamil vowel within the vowel chart, we can classify vowels; Figure 2 shows the position of the Tamil vowels for the male speakers from the data obtained as shown from the Table I and its vowel triangle representation.

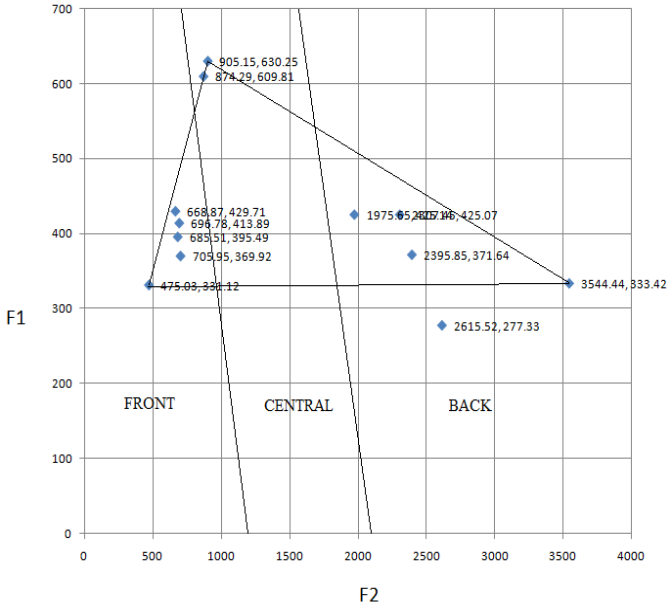

Fig. 2. Tamil vowel diagram for the male speakers

Figure 3 shows the position of the Tamil vowels for the female speakers from the data obtained as shown from the Table II and its vowel triangle representation.

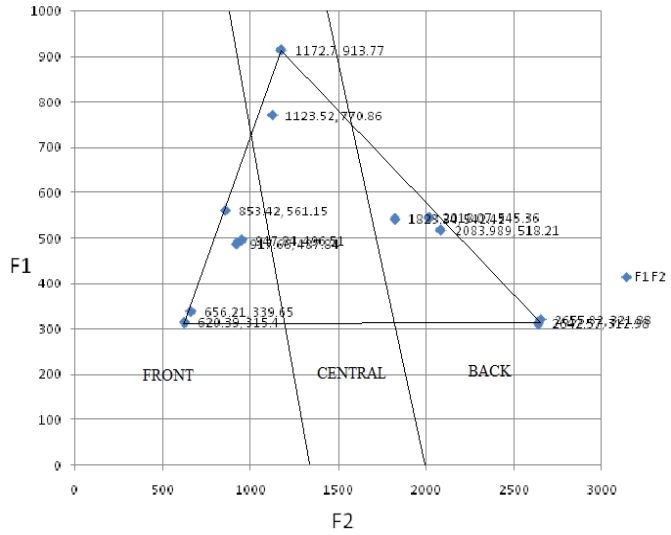

Fig. 3. Tamil vowel diagram for the female speakers

Figure 4 shows the position of the Tamil vowels for both male and female speakers from the data obtained as shown from the Table III and its vowel triangle representation.

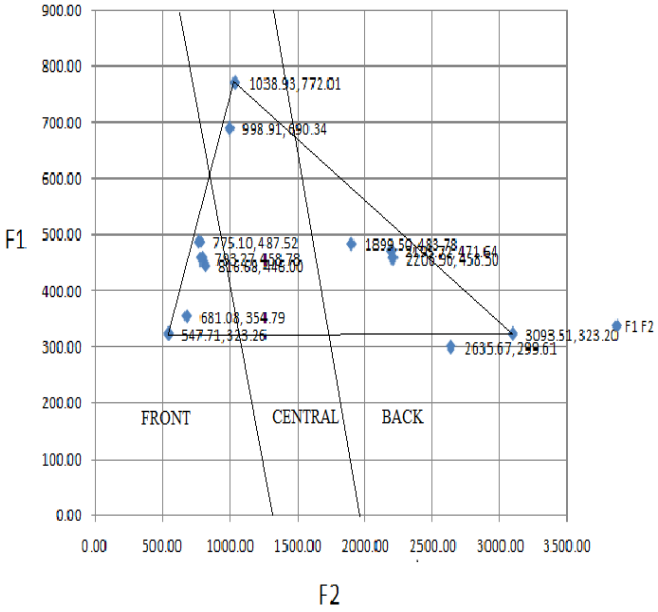

Fig. 4. Tamil vowel diagram for the both male and female speakers combined

Based on above diagram, Tamil vowels are classified as Front, central and back with respect to their position on the vowel chart, and the results are tabulated.

VI. RESULTS AND CLASSIFICATION OF TAMIL VOWELS

Vowels are unique in that their main features do not contain deferent in voicing or articulators, instead it depend upon tongue position when voiced, especially with the position of tongue forwardness or backwardness with respect to the vowel produced. The Tamil vowel diagram help us to identify how far forward the tongue is located in the oral cavity during Tamil vowel production, there by classifying Tamil vowel with respect to tongue front-back position.

TABLE IV. Classification of Tamil vowel for male speakers

Format Frequency		**FRONT**				
		ஐ	எ	ஏ	இ	ஈ
F1 Hz	Average	425.14	425.07	371.64	333.42	277.33
F2 Hz	Average	1975.65	2307.45	2395.85	3544.44	2615.52
Format Frequency		**CENTRAL**				
		அ	ஆ			
F1 Hz	Average	609.81	630.25			
F2 Hz	Average	874.29	905.15			
Format Frequency		**BACK**				
		உ	ஊ	ஒ	ஓ	ஔ
F1 Hz	Average	331.12	369.92	395.49	429.71	413.89
F2 Hz	Average	475.03	705.95	685.51	668.87	696.78

TABLE V. Classification of Tamil vowel for female speakers

Format Frequency		FRONT				
		ஐ	எ	ஏ	இ	ஈ
F1 Hz	Mean	542.42	518.12	545.36	312.98	321.88
F2 Hz	Mean	1823.34	2083.98	2018.07	2642.57	2655.82
Format Frequency		**CENTRAL**				
		அ	ஆ			
F1 Hz	Mean	770.86	913.77			
F2 Hz	Mean	1123.52	1172.7			
Format Frequency		**BACK**				
		உ	ஊ	ஒ	ஓ	ஔ
F1 Hz	Mean	315.4	339.65	496.51	487.84	561.15
F2 Hz	Mean	620.39	656.21	947.84	917.66	853.42

TABLE VI. Classification of Tamil vowel for both male and female speakers combined

Format Frequency		FRONT				
		ஐ	எ	ஏ	இ	ஈ
F1 Hz	Mean	483.78	471.64	458.5	323.2	299.61
F2 Hz	Mean	1899.5	2195.72	2206.96	3093.51	2635.67
Format Frequency		**CENTRAL**				
		அ	ஆ			
F1 Hz	Mean	690.34	772.01			
F2 Hz	Mean	998.91	1038.93			
Format Frequency		**BACK**				
		உ	ஊ	ஒ	ஓ	ஔ
F1 Hz	Mean	323.26	354.79	446	458.78	487.52
F2 Hz	Mean	547.71	681.08	816.68	793.27	775.1

Based on the above Tamil vowels classification it is very clear that whoever the speaker is the Tamil vowels classification remains the same.

TABLE VII. Tamil vowels classification

S.no	TONGUE FRONT-BACK POSITION	VOWELS CLASSIFICATION
1	FRONT	எ,ஏ,இ,ஈ,ஐ
2	CENTRAL	அ,ஆ
3	BACK	உ,ஊ,ஒ,ஓ,ஔ

TABLE VII shows the classification of Tamil vowels with respect to the tongue front-back position based on Tamil vowel diagram.

VII. CONCLUSION

Taking into consideration the fact, that different language has its own special linguistic particularities, so in order to design a proper speech recognition system we need to understand the phonetic characteristics of the particular language being used for the speech recognition system, Tamil language has its own special linguistic particularities, if we want to design a Tamil speech recognition system we need to understand its phonetic characteristics.

Tamil vowels are long in duration as compared to Tamil consonant sounds and are spectrally well defined. As such they are usually easily recognized and therefore contribute significantly to our ability to recognize speech. Therefore classification of Tamil vowels with respect to its phonetic characteristics plays the major role in Tamil speech recognition, since vowels are unique in that their main features do not contain deferent in voicing or articulators, instead it depend upon tongue position when voiced, especially with the position of tongue forwardness or backwardness with respect to the vowel produced. Therefore the classification of Tamil vowels with respect to the tongue front-back position based on Tamil vowel diagram not only classified the Tamil vowels but also provided with vital information about the spectral characteristics of each and every Tamil vowels. This can provide useful information for the modeling of Tamil speech recognition system.

REFERENCES

[1] http://en.wikipedia.org/wiki/Tamil_language.
[2] Annamalai, E.; Steever, S.B. (1998), "Modern Tamil", in Steever, Sanford, The Dravidian Languages, London: Routledge, pp. 100–128, ISBN 0-415-10023-2
[3] Clark, John; & Yallop, Colin. (1995). An introduction to phonetics and phonology (2nd ed.). Oxford: Blackwell. ISBN 0-631-19452-5.
[4] O'Grady, William, et al. (2005). Contemporary Linguistics: An Introduction (5th ed.). Bedford/St. Martin's. ISBN 0-312-41936-8.
[5] Lawrence Rabiner, Biing-Hwang Jung, "Fundamentals of Speech Recognition", PEARSON EDUCATION (Singapore) PRIVATE LIMITED, Indian Branch, 482 F.I.E Patpargans, Delhi 110092, India, © 1993.
[6] Paul Skandera and Peter Burleigh (2005). A Manual of English Phonetics and Phonology. Gunter Narr Verlag. pp. 33–34. ISBN 3823361252.
[7] Bickford, Anita (2006). Articulatory Phonetics: Tools For Analyzing The World's Languages (4th ed.). Summer Institute of Linguistics. ISBN1-55671-165-4.