Using Exact Formant Structure of Persian Vowels as a Cue for Forensic Speaker Recognition

Mojtaba Namvar Fargi
Shahla Sharifi
Mohammad Reza Pahlavan-Nezhad
Azam Estaji
Mehi Meshkat Aldini

Linguistics Department, Ferdowsi University of Mashhad
Mashhad, Iran

==
Language in India www.languageinindia.com **ISSN 1930-2940 Vol. 13:6 June 2013**
==

Abstract

Forensic phonetics is subfield of forensic linguistics in which using acoustic information and phonetic features s are investigated for completing forensic cases in which one of the existing evidences is a voice record related to the guilty. One of the most important tasks of forensic phoneticians is forensic speaker recognition. For doing this the phonetician is asked to estimate the degree of similarity between the given records of the guilt's speech and the suspected(s) and determine that whether these two sound evidences match to each other or not.

The objectives of this study which is conducted on the sound data from 10 Persian native speakers of both sexes, is to investigate the possibility of using exact formant structure of vowels as a cue for forensic speaker recognition tasks.

The results of this study show that using vowel space and exact formant structure of vowels may be a useful means with perfect reliability in tasks related to forensic speaker recognition.

Keywords: *forensic linguistics, forensic phonetics, speaker recognition, acoustic phonetics*.

1. Introduction

Forensic linguistics, deals with the topics such as Forensic Language (Tiersma, 1999), Forensic Semantics (Langford, 2000: 72-94), Forensic Discourse (Bavelas, & Gibson, 1994: 189-206), False testimony (Harris, 2001: 53-74) and forensic phonetics (Aqagolzadeh, 1391).

Forensic Phonetics, which is a subfield of forensic linguistics, deals with issues such as forensic speaker recognition, forensic speech recognition, forensic speaker identification and forensic speaker verification and determining the truth or falseness of phone calls (Hollein, 1990: 190-191).

Forensic phonetics tries to determine the probability that a recorded voice or recorded phone call which is attributed to the accused person can be the same as the voice of a suspect.

This probability can help a court or police or related organization to solve a judicial case. It may be also useful in some more complex cases in which it's tried to determine whether a speaker had stress while doing the crime or not and to get some information about his/her emotional status out of the recorded voice.

In this study, we used Persian vowel space and exact formant structure of Persian vowels for forensic speaker recognition.

The number of juridical cases in which courts require speaker recognition through recorded voices, is increasing. In such cases the phonetician is asked to do forensic speaker recognition. In other words, he is asked to make a comparison between the recorded voice of the guilty while doing the crime and the suspect's to help confirming the guilt or exonerating the suspect doing this (Nolan, McDougall, Jong & Hudson, 2006).

The task of speaker recognition is affected by numerous varieties. In particular, speakers change their voices under conditions like their familiarity with the addressee, emotional statuses, the degree of formality of the situation, the degree of background noise and such (Nolan, 1997: 748).Also the person's voice changes with changes in the health status and that the speaker may change his voice deliberately or copy another person's voice (Nolan, McDougall, Jong & Hudson, 2006).

In forensic speaker recognition, multiple variables must be paid attention to. For instance, Hollein (1990) states that these variables include: non-consecutiveness of the records, vocal changes due to the recording system used, and the varieties being formed in the speech of a single person for many reasons such as changes in emotional or health statuses and even the intentional voice change (voice imitation) (Hollein, 1990: 190-191) .

Kinoshita (1998), investigates how stylistic features affect the sound features of Japanese speakers WRT forensic phonetics (Kinoshita, 1998).

In the 1960s, the scientists tried to use the speech spectrogram apparatuses as a means for speaker's recognition. (Kersta, 1962, Stevens, Williams, Carbonelli and Woods, 1968, Over, Lashbrook, Pedrev, Nicol and Nash,1972). However, the findings were limited, with the analysts working with these tools. At that time, the computer technology was not that advanced to enable the experts in doing so. Advancements in the computer technology after the 1960s, made a new wave of studies in the field of speaker recognition. Although these studies made improvements to the text-based speaker recognition systems, they still had considerable deficiencies in the field of text-independent speaker recognition which accompanied varieties like channel and the speaker and these studies were not successful as expected (Rodman, 1997).

Despite so many advancements in today's systems of speaker recognition, there still remain shortcomings which delay the recognition task or reduce the efficiency and the accuracy of the task; so that the recognition accuracy is affected by the environmental changes, speaker variation and the natural linguistic input conversely (Herbig, Gerl, Minker, 2011).

2. **Data collection and methods**

The data of this study, included 100 hundred minutes monologue and speech of 10 Persian speakers 5 of which were male and the other 5 were female who were in the age group of 20 to 35. These people's voices were recorded using professional voice recording devices in stereo format at sample rate 44100 and the Bit Depth 16 in Sorush-e-Sima Recording Studio of Khorasan Razavi Broadcasting Center. The software recording studio was Adobe Audition.

Selecting the people whose voices are studied is done randomly and there were only three factors interfering with it. First thing is that they be in the age group of 20 to 35 to make sure their voice is different with that of children and adolescents and also with that of the old. The second is that among these, 5 speakers are male and the other 5 ones are females and the third factor was that these people enjoy with the standard dialect or the dialect close to the standard Persian.

All voices were recorded at once. Each speaker was asked to read a pre-determined text with the typical and usual routine tone. This text included Persian language vowels separately, one-syllabic words having one consonant and one vowel for all the vowels and consonants of the Persian language, 140 words each of which was repeated twice randomly and 34 sentences in a conversation form, which included all of the separate words from the text. The used words in this investigation were selected according to the table Samareh (1381) had provided for the allomorphs of the Persian language morphs, and it was tried to embed in these data all the different contexts in which consonants and vowels of Persian language might be used.

The speakers were asked to make a pause in between the vowels, syllables and also the words while pronouncing them to ease the separation of the sound data related to these voices, syllables and the words and to make sure that the data have not merged with each other. In the part related to the reading the conversational sentences the speakers were asked to read those texts just like a typical speech and before the recording began, they had already read the text for an hour to make sure that no problems would pop up while reading the text and that the natural velocity and the speech style is assured while reading.

3. **Data Analysis**

The prime hypothesis of this key is achieved from the evident relation between the formants and the sizes of vocal tract. The formants are under direct effect of the shape of the articulation system as the resonance-maker device and the resonance of each individual's voice differs with another one and it is this very difference in the people's resonances that enables us to recognize people through their voices. This primary hypothesis is studied with 18 various experiments in this thesis.

In the first two experiments, the patterns of exact place of the formants is achieved using the SPSS software to use the frequency mean for the formants of vowels for 5 male speakers using graphs that their horizontal axis is F2 and the vertical axis is F. And the 5 provided patterns were put into one graph to find out the similarity or difference of them. Then the same process was applied to the vowels of 5 female speakers.

In the third to the twelfth experiments, the achieved patterns from the structure of vowels of each speaker were mapped in separate graphs using F1 & F2 in five sets of different words that the words in each set were also randomly different with the words from the other sets and the patterns related to each speaker were put into one single graph in five sets of different words which were marked with different colors to distinguish between the degree of their similarity or difference

These twelve experiments were reiterated once more and this time F1 and the space between F1 & F2 was used for F1 & F2 or the F1-F2 and the results were also studied.

The point which is present here is that in these surveys for more assurance of the results, various words and different voice contexts were used for the extraction of formants of vowels while such variety is not needed and the accused could be asked to repeat the same words and sentences which exist in the pre-existing data in different times to extract the formants of the vowels in the same voice contexts and the same places. But using the variety in the words and random selection of them brought more trust to the achieved results.

In the figure below, the first formant is marked by F1, the second formant is marked by F2 and the third formant is marked by F3.

The resulted frequency is also noted below each of these variables. The red points are the points that the Praat Software has presented which of course they are sometimes of low accuracy and here the formants are manually presented with more accuracy with making the graph bigger or smaller.

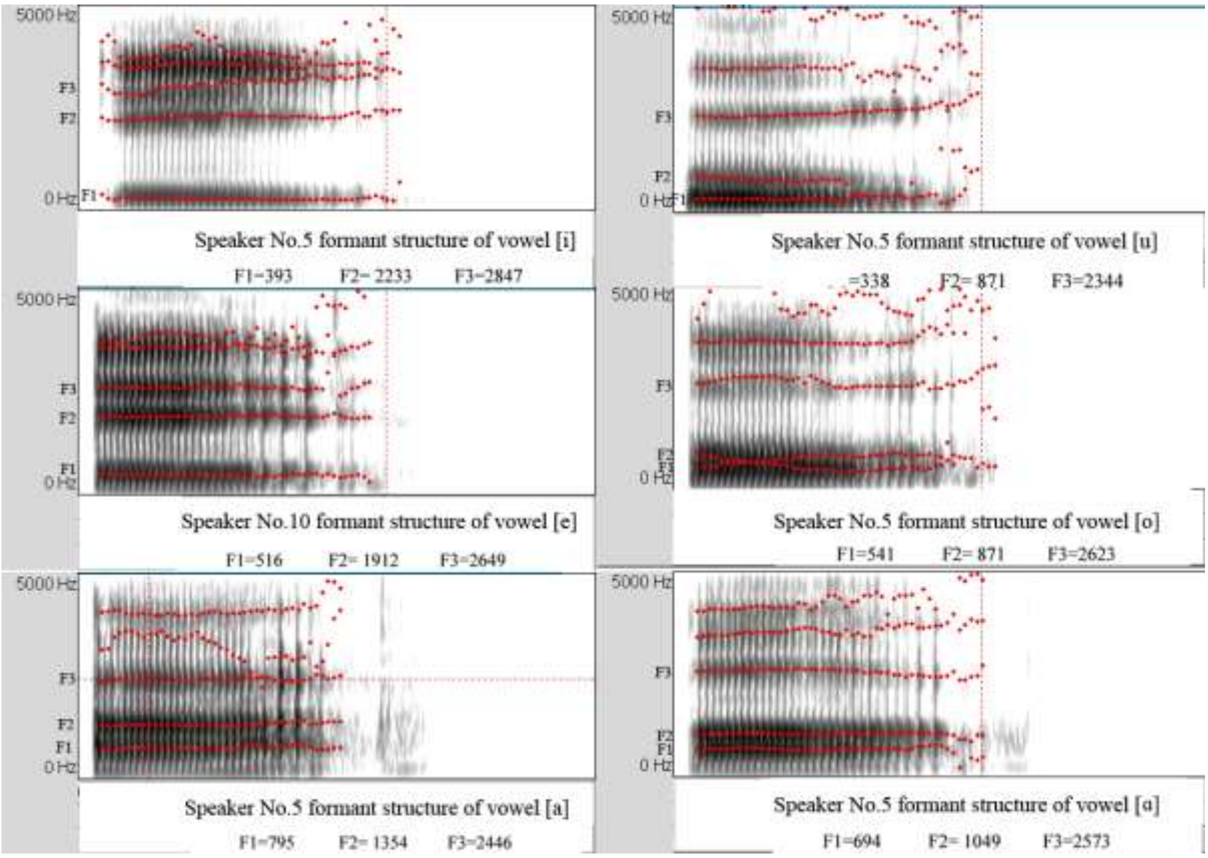

Figure1: Formation of F1, F2 and F3 in pronouncing vowels of the Persian language by Speaker No. 5

In this figure, each of the 6 Persian language vowels, are presented in the sound data of a male speaker (Speaker No. 5). Their arrangements in this figure are based up on the same traditional pattern of displaying the vowels of the Persian language, in a way that the posterior vowels are at the right side and the anterior ones are seen at the left side and the up and down figures are in row the highest and the lowest vowels and the middle figures indicate the middle vowels.

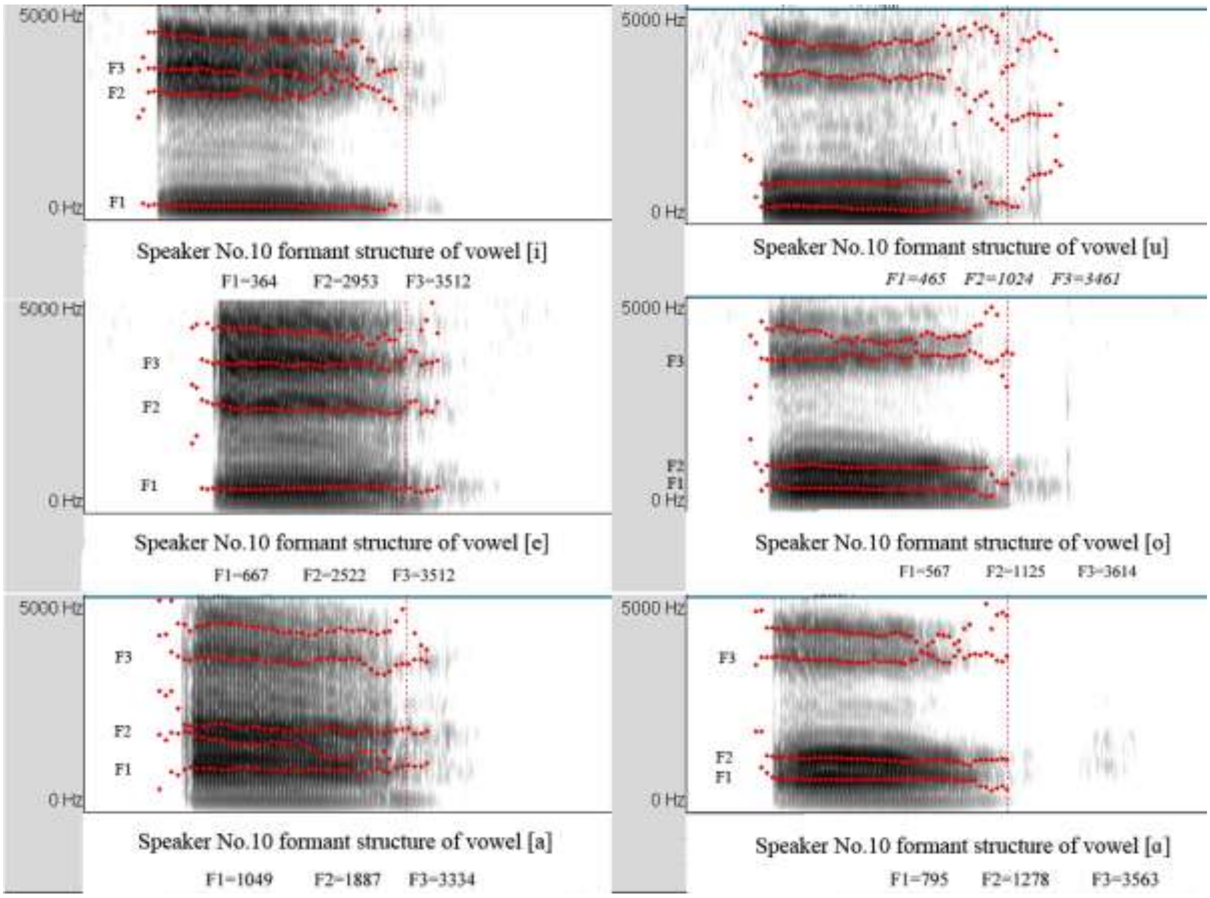

Figure 2: Formation of F1, F2 and F3 in pronouncing vowels of the Persian language by Speaker No. 10

In this figure, each of the 6 Persian language vowels, are presented in the sound data of a female speaker (Speaker No. 10). Their arrangements in this figure are based up on the same traditional pattern of displaying the vowels of the Persian language, in a way that the posterior vowels are at the right side and the anterior ones are seen at the left side and the up and down figures are in row the highest and the lowest vowels and the middle figures indicate the middle vowels.

Speaker	Vowel	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
Speaker No. 1	Vowel [ɩ]	338	1887	3182	1549	1295
	Vowel [ɛ]	516	1760	2598	1244	838
	Vowel [ɑ]	744	1532	2293	788	761
	Vowel [ʋ]	415	846	2169	431	1323
	Vowel [o]	491	998	2217	507	1219
	Vowel [A]	592	998	2471	406	1473
Speaker No. 2	Vowel [ɩ]	338	2471	2877	2133	406
	Vowel [ɛ]	465	1862	2623	1397	761
	Vowel [ɑ]	821	1532	2446	711	914
	Vowel [ʋ]	313	795	2547	482	1752
	Vowel [o]	465	770	2547	305	1777
	Vowel [A]	693	1024	2801	331	1777
Speaker No. 3	Vowel [ɩ]	262	2446	3182	2184	736
	Vowel [ɛ]	440	2090	2750	1650	660
	Vowel [ɑ]	948	1582	2725	634	1143
	Vowel [ʋ]	364	846	2598	482	1752
	Vowel [o]	416	897	2573	481	1676
	Vowel [A]	684	1103	2911	419	1808

Speaker No. 4	Vowel [ɩ]	288	2243	3055	1955	812
	Vowel [ɛ]	491	2039	2674	1548	635
	Vowel [α]	846	1608	2892	762	1285
	Vowel [ʊ]	267	735	2446	468	1711
	Vowel [o]	491	871	2496	380	1625
	Vowel [A]	694	1024	2598	330	1574
Speaker No. 5	Vowel [ɩ]	393	2233	2847	1840	614
	Vowel [ɛ]	516	1912	2649	1396	737
	Vowel [α]	795	1354	2446	559	1092
	Vowel [ʊ]	338	871	2344	533	1473
	Vowel [o]	541	871	2623	330	1752
	Vowel [A]	694	1049	2573	355	1524

Table 1: The frequency of formants of the vowels pronounced by speakers 1 to 5 (Male Speakers)

Speaker	Vowel	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
Speaker No. 6	Vowel [ɩ]	440	2700	3360	2260	660
	Vowel [ɛ]	577	2344	2877	1767	533
	Vowel [α]	1084	1661	2776	577	1115
	Vowel [ʊ]	415	897	2750	482	1853
	Vowel [o]	516	998	2826	482	1828
	Vowel [A]	846	1176	2623	330	1447
Speaker No. 7	Vowel [ɩ]	363	3182	3537	2819	355
	Vowel [ɛ]	592	2649	3334	2057	685
	Vowel [α]	1151	1989	3004	838	1015
	Vowel [ʊ]	415	897	2979	482	2082
	Vowel [o]	465	973	3157	508	2184
	Vowel [A]	845	1265	3105	420	1840
Speaker No. 8	Vowel [ɩ]	312	2953	3360	2641	407
	Vowel [ɛ]	516	2496	3157	1980	661
	Vowel [α]	1075	1963	3334	888	371
	Vowel [ʊ]	313	846	2776	533	1930
	Vowel [o]	491	948	2826	457	1878
	Vowel [A]	795	1227	3131	432	1904
Speaker No. 9	Vowel [ɩ]	338	2446	3588	2108	1142
	Vowel [ɛ]	668	2243	3055	1575	812
	Vowel [α]	1024	1862	2852	838	990
	Vowel [ʊ]	364	846	2891	482	2045
	Vowel [o]	516	896	3004	380	2108
	Vowel [A]	872	1176	2928	304	1752
Speaker No. 10	Vowel [ɩ]	364	2953	3512	2589	559
	Vowel [ɛ]	667	2522	3512	1855	990
	Vowel [α]	1049	1887	3334	838	1447
	Vowel [ʊ]	465	1024	3461	559	2437
	Vowel [o]	576	1125	3614	558	2489
	Vowel [A]	795	1278	3563	483	2285

Table 2: The frequency of formants of the vowels pronounced by speakers 6 to 10 (Female Speakers)

This table is arranged for calculating the average of formants frequency for each vowel in table below:

Vowel	Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
Vowel [ɩ]	**Speaker No. 1**	338	1887	3182	1549	1295
	Speaker No. 2	338	2471	2877	2133	406
	Speaker No. 3	262	2446	3182	2184	736
	Speaker No. 4	288	2243	3055	1955	812
	Speaker No. 5	393	2233	2847	1840	614
The average frequency of formants in the vowel [ɩ] pronounced by male speakers		323.8	2256	3028.6	1932.2	772.6
	Speaker No. 1	516	1760	2598	1244	838
	Speaker No. 2	465	1862	2623	1397	761

Vowel [ɛ]	Speaker No. 3	440	2090	2750	1650	660
	Speaker No. 4	491	2039	2674	1548	635
	Speaker No. 5	516	1912	2649	1396	737
The average frequency of formants in the vowel [ɛ] pronounced by male speakers		485.6	1932.6	2658.8	1447	726.2
vowel [ɑ]	Speaker No. 1	744	1532	2293	788	761
	Speaker No. 2	821	1532	2446	711	914
	Speaker No. 3	948	1582	2725	634	1143
	Speaker No. 4	846	1608	2893	762	2285
	Speaker No. 5	795	1354	2446	559	1092
The average frequency of formants in the vowel [ɑ] pronounced by male speakers		830.8	1521.6	2560.6	690.8	1239

Table3: The frequency of the formants of vowels [ɩ] ،[ɛ] و [ɑ] in the sound data related to the speakers 1 to 5 (male)

Vowel	Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
Vowel [ʊ]	Speaker No. 1	415	846	2169	431	1323
	Speaker No. 2	313	795	2547	482	1752
	Speaker No. 3	364	846	2598	482	1752
	Speaker No. 4	267	735	2446	468	1711
	Speaker No. 5	338	871	2344	533	1473
The average frequency of formants in the vowel [ʊ] pronounced by male speakers		339.4	818.6	2420.8	479.2	1602.2
Vowel [o]	Speaker No. 1	491	998	2217	507	1219
	Speaker No. 2	465	770	2547	305	1777
	Speaker No. 3	416	897	2573	481	1676
	Speaker No. 4	491	871	2496	380	1625
	Speaker No. 5	541	871	2623	330	1752
The average frequency of formants in the vowel [o] pronounced by male speakers		480.8	881.4	2491.2	400.6	1609.8
Vowel [A]	Speaker No. 1	592	998	2471	406	1473
	Speaker No. 2	693	1024	2801	331	1777
	Speaker No. 3	684	1103	2911	419	1808
	Speaker No. 4	491	871	2496	380	1625
	Speaker No. 5	694	1049	2573	355	1524
The average frequency of formants in the vowel [A] pronounced by male speakers		630.8	1009	2650.4	378.2	1641.4

Table 4 continues: The frequency of the formants of vowels [ʊ] ،[o] و [A]in the sound data related to the speakers 1 to 5 (male speakers)

The frequency of the first, second and the third formants is calculated for the speakers 6 to 10 (Female Speakers) and its results are noted here in a nutshell.

Vowel	Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
vowel [ɩ]	Speaker No. 6	440	2700	3360	2260	660
	Speaker No. 7	363	3182	3537	2819	355
	Speaker No. 8	312	2953	3360	2641	407
	Speaker No. 9	338	2446	3588	2108	1142
	Speaker No. 10	364	2953	3512	2589	559
The average frequency of formants in the vowel [ɩ] pronounced by female speakers		363.4	2846.8	3471	2483	624.6
vowel [ɛ]	Speaker No. 6	577	2344	2877	1767	533
	Speaker No. 7	592	2649	3334	2057	685
	Speaker No. 8	516	2496	3157	1980	661
	Speaker No. 9	668	2243	3055	1575	812
	Speaker No. 10	667	2522	3512	1855	990
The average frequency of formants in		604	2450.8	3187	1847	736.2

the vowel [ɛ] pronounced by female speakers						
vowel [ɑ]	Speaker No. 6	1084	1661	2776	577	1115
	Speaker No. 7	1151	1989	3004	838	1015
	Speaker No. 8	1075	1963	3334	888	371
	Speaker No. 9	1024	1862	2852	838	990
	Speaker No. 10	1049	1887	3334	838	1447

Table 5: The frequency of the formants of vowels [ɩ] ، [ɛ] و [ɑ] in the sound data related to the speakers 6 to 10 (female speakers)

Vowel	Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
vowel[ʋ]	Speaker No. 6	415	897	2750	482	1853
	Speaker No. 7	415	897	2979	482	2082
	Speaker No. 8	313	846	2776	533	1930
	Speaker No. 9	364	846	2891	482	2045
	Speaker No. 10	465	1024	3461	559	2437
The average frequency of formants in the vowel [ʋ] pronounced by female speakers		394.4	902	2971	507.6	2069.4
vowel [o]	Speaker No. 6	516	998	2826	482	1828
	Speaker No. 7	465	973	3157	508	2184
	Speaker No. 8	491	948	2826	457	1878
	Speaker No. 9	516	896	3004	380	2108
	Speaker No. 10	576	1125	3614	558	2489
The average frequency of formants in the vowel [o] pronounced by female speakers		512.8	988	3085	477	2097.4
vowel[A]	Speaker No. 6	846	1176	2623	330	1447
	Speaker No. 7	845	1265	3105	420	1840
	Speaker No. 8	795	1227	3131	432	1904
	Speaker No. 9	872	1176	2928	304	1752
	Speaker No. 10	795	1278	3563	483	2285
The average frequency of formants in the vowel [A] pronounced by female speakers		830.6	1224.4	3070	393.8	1845.6

Table 6 continues: The frequency of the formants of vowels [ʋ] ، [o] و [A]in the sound data related to the speakers 6 to 10 (female speakers)

The summary of the above tables is prepared at the two below tables:

Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
The average frequency of formants in the vowel [ɩ] pronounced by male speakers	323.8	2256	3028.6	1932.2	772.6
The average frequency of formants in the vowel [ɛ] pronounced by male speakers	485.6	1932.6	2658.8	1447	726.2
The average frequency of formants in the vowel [ɑ] pronounced by male speakers	830.8	1521.6	2560.6	690.8	1239
The average frequency of formants in the vowel [ʋ] pronounced by male speakers	339.4	818.6	2420.8	479.2	1602.2
The average frequency of formants in the vowel [o] pronounced by male speakers	480.8	881.4	2491.2	400.6	1609.8
The average frequency of formants in the vowel [A] pronounced by male speakers	630.8	1009	2650.4	378.2	1641.4

Table7: The frequency Average of the formants of each vowel in the sound data related to the speakers 1 to 5 (male speakers)

Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
The average frequency of formants in the vowel [ɩ] pronounced by female speakers	363.4	2846.8	3471	2483	624.6
The average frequency of formants in the vowel [ɛ] pronounced by female speakers	604	2450.8	3187	1847	736.2
The average frequency of formants in the vowel [ɑ] pronounced by female speakers	1076.6	1872.4	3060	795.8	987.6
The average frequency of formants in the vowel [ʋ] pronounced by female speakers	394.4	902	2971	507.6	2069.4
The average frequency of formants in the	512.8	988	3085	477	2097.4

vowel [o] pronounced by female speakers					
The average frequency of formants in the vowel [A] pronounced by female speakers	830.6	1224.4	3070	393.8	1845.6

Table 8: The frequency Average of the formants of each vowel in the sound data related to the speakers 6 to 10 (female speakers)

In the following tables, the frequency average of the formants for the sum of speakers 1to10 is tallied.

Speaker	F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
The average frequency of formants in the vowel [ɩ] pronounced for the sum of speakers 1-10	343.6	2551.4	3249.8	2207.6	698.6
The average frequency of formants in the vowel [ɛ] pronounced for the sum of speakers 1-10	544.8	2191.7	2922.9	1647	731.2
The average frequency of formants in the vowel [ɑ] pronounced for the sum of speakers 1-10	953.7	1697	2810.3	743.3	1113.3
The average frequency of formants in the vowel [ʋ] pronounced for the sum of speakers 1-10	366.9	860.3	2695.9	493.4	1835.8
The average frequency of formants in the vowel [o] pronounced for the sum of speakers 1-10	496.8	934.7	2788.1	438.8	1853.6
The average frequency of formants in the vowel [A] pronounced for the sum of speakers 1-10	730.7	1116.7	2860.2	386	1743.5

Table 9: The frequency Average of the formants of each vowel in the sound data related to the sum of speakers 1 to 10

The most prominent similarity of the posterior vowels is the close connection of F1 and F2 in them. As it is observed in the figures, highness of the vowel has a converse relationship with the degree of its F1. In other words, the highest the vowel, the less is the frequency of its F1. F3 is more related to the roundedness feature. Of course it doesn't maintain a direct relation with roundedness.

According to the relation of each formant with the form of speech system, these relationships among formants and the features of vowels are true:

F1 has a converse relation with the vowel degree of highness or the degree of openness or closeness of the mouth. When the tongue moves from a higher place to a lower one, the space of the mouth increase and the space of pharyngeal cavity decreases and since the space of the pharyngeal cavity has a converse relation with F1, with bringing the tongue down, F1 increases and with its going up, it decreases.

Regarding the openness or closeness degree of the mouth, it can also be stated that the more the mouth goes toward closeness F1 decreases and the more the mouth opens, F1 increases.

F2 has a direct relation with the changes of mouth cavity. Posterior vowels have a higher F2 and anterior ones have a lower F2.

In sum, the space between F1 & F2 is closer in anterior vowels and farther in the posterior ones.

The data of this study is compatible with the relations that Kent et al. (1996) suggested for the relation among F1, F2 & F3 with the openness or closeness of the vowels. These relations are as follows:

The degree of being anterior or posterior	The relation of the space F1، F2 و F3
Anterior < Posterior vowels vowels	**The degree of F2-F1**
Anterior < Posterior vowels vowels	**The degree of F3-F2**

Table 10: The relation between F1, F2 and F3 with being anterior or posterior of vowels (Kent et al, 1996)

2-1-4 The Persian Language Vowel Space

In this study, regarding the achieved average for the formants of vowels, the Persian language vowel space is first calculated separately for the male and female speakers and then using the calculation of the average formants of vowels of both male and female speakers it is shown in the graphs below.

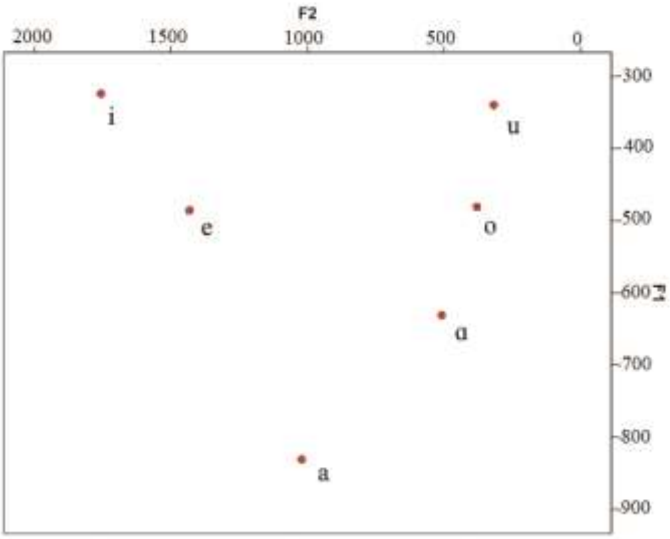

Figure 3: The achieved vowel space from the average first and second formants of the pronounced vowels by the speakers 1- 5 (male speakers)

The vertical axis indicates the first formant and the horizontal axis indicates the second formant.

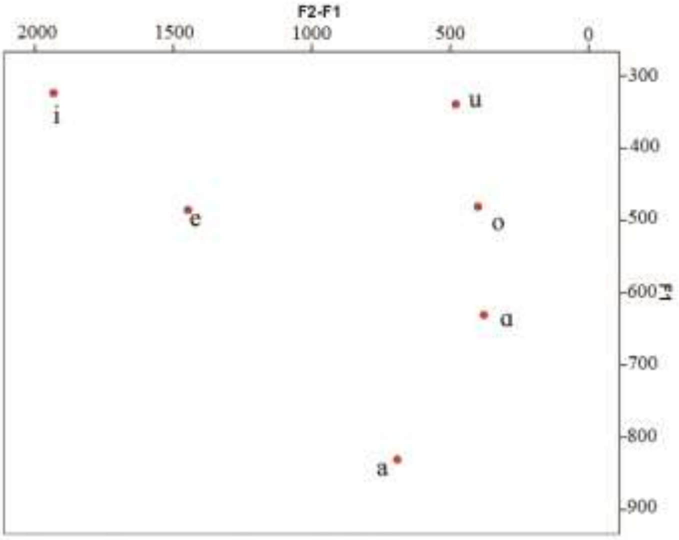

Figure 4: The achieved vowel space from the average first formant and the distance between the first and the second formants of the pronounced vowels by the speakers 1- 5 (male speakers)

The vertical axis indicates the first formant and the horizontal axis indicates distance between the first and the second formants

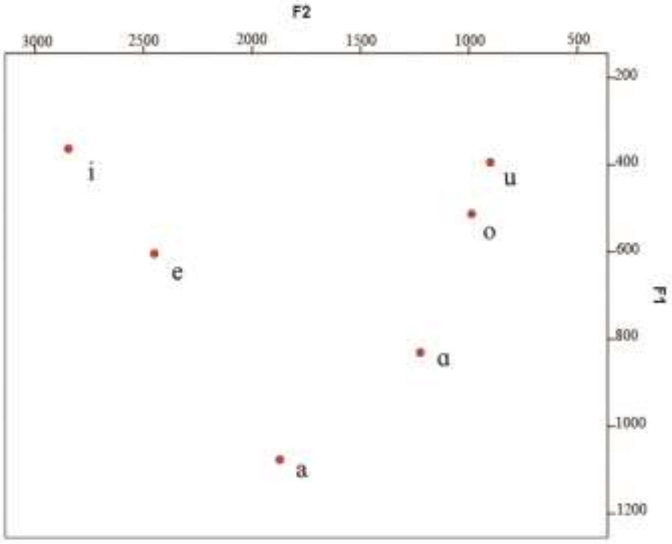

Figure 5: The achieved vowel space from the average first and second formants of the pronounced vowels by the speakers 6- 10 (female speakers)

The vertical axis indicates the first formant and the horizontal axis indicates the second formant.

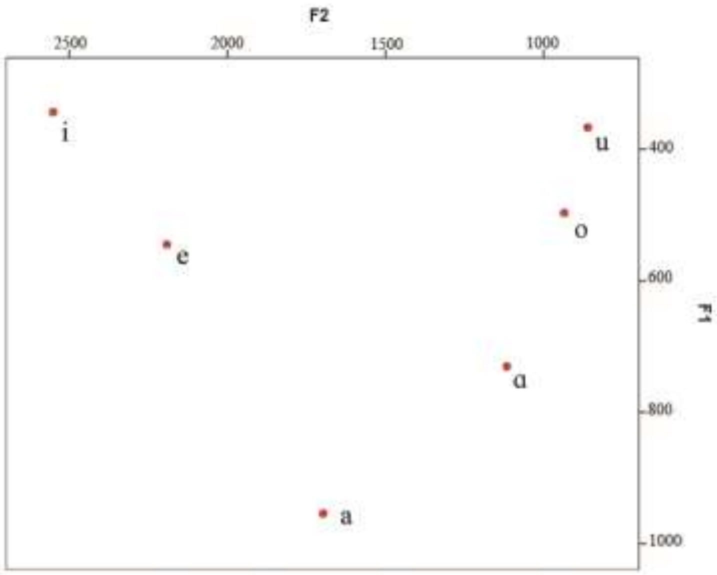

Figure 6: The achieved vowel space from the average first and second formants of the pronounced vowels by the speakers 1- 10 (5 males speakers and 5 female speakers)

The vertical axis indicates the first formant and the horizontal axis indicates the second formant.

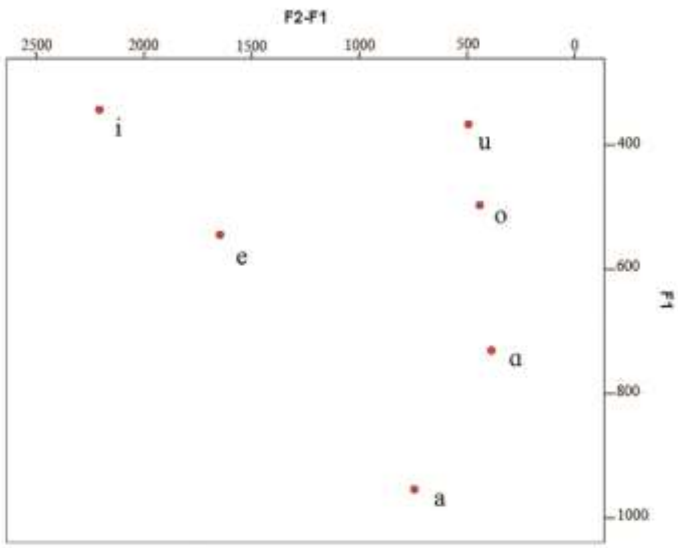

Figure 6: The achieved vowel space from the average first formant and the distance between the first and the second formants of the pronounced vowels by the speakers 1- 10 (5 males speakers and 5 female speakers)

The vertical axis indicates the first formant and the horizontal axis indicates distance between the first and the second formants

3. The formant formation of vowels as a speaker-dependent feature

Since the formants frequency is dependent to the general shape of the speech organ and different people have different features of jaws and the internal sizes of speech organs, this hypothesis was created in this study that it is possible that a vowel space in each person have a unique behavior and could be used as a key for the forensic speaker recognition.

For this purpose, for each person regarding the achieved formants for the vowels, the vowel space was once mapped regarding the F1& F2 and a second time with regarding the space between F1, F2. Then this vowel space was delineated.

For 5 male speakers the vowel spaces were mapped in one single graph and the vowel space relevant to each speaker was marked a different color. The achieved result was so substantial. None of the vowel spaces were in accordance with each other and they had obvious differences with each other.

This feature was there for both graphs. In each graphs, the vowel spaces of different people had different features with each other.

The figures related to this experiment are shown below.

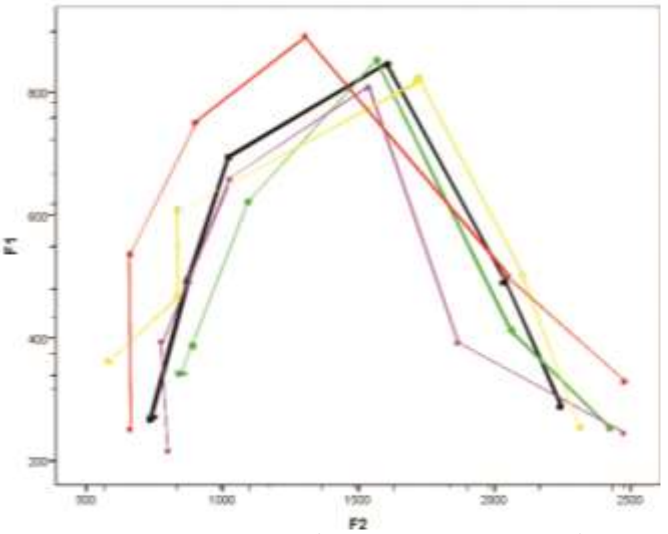

Figure 7: The graph related to the exact structure of vowels according to F1, F2 for the speakers1-5 (male)

The vertical axis indicates the first formant and the horizontal axis indicates the second formant.

The color yellow, indicates the place of the first and second formants of the vowels of the speaker no. 1 which are connected to each other with yellow lines.

The color violet indicates the place of the first and second formants of the vowels of the speaker no. 2 which are connected to each other with violet lines.

The color green indicates the place of the first and second formants of the vowels of the speaker no. 3 which are connected to each other with green lines.

The color black, indicates the place of the first and second formants of the vowels of the speaker no. 4 which are connected to each other with black lines.

The color red indicates the place of the first and second formants of the vowels of the speaker no. 5 which are connected to each other with red lines.

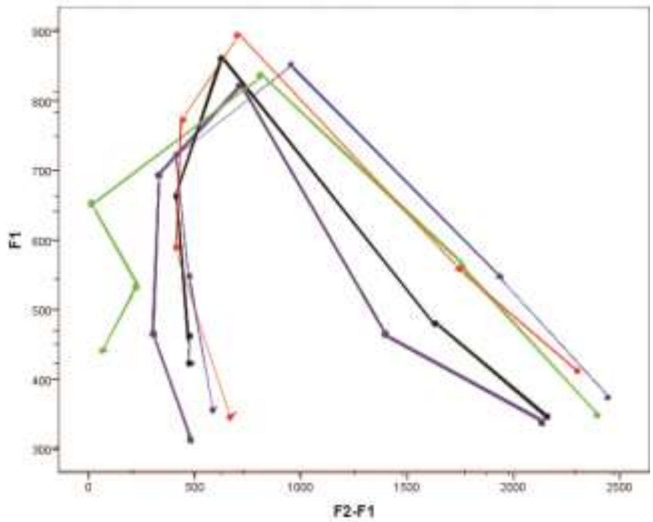

Figure 8: The exact vowel structure according to F2 and the distance between F1 & F2 for the speakers1-5 (male)

The vertical axis indicates the first formant and the horizontal axis indicates the distance between the first and the second formant.

The color green indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 1 which are connected to each other with green lines.

The color violet indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 2 which are connected to each other with violet lines.

The color black indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 3 which are connected to each other with black lines.

The color blue indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 4 which are connected to each other with blue lines.

The color yellow indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 5 which are connected to each other with yellow lines.

As it is observed in these two graphs, the exact structure of the formants of vowels for all speakers is incompatible against each other and the exact values of these structures are different for each speaker with another one. This difference solely suffices to be used in forensic speaker recognition. This was also done for the 5 female speakers and had the same results which can be observed in the following graphs.

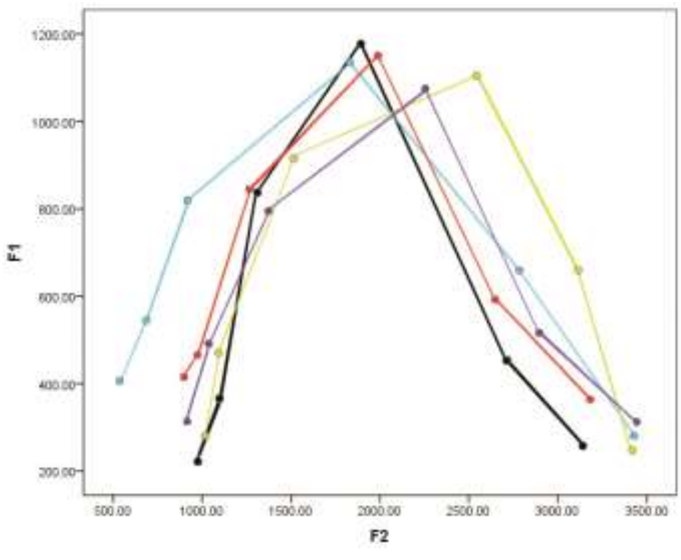

Figure 9: The exact structure of vowels according to F1, F2 for the speakers6-10 (female)
The vertical axis indicates the first formant and the horizontal axis indicates the second formant.

The color black, indicates the place of the first and second formants of the vowels of the speaker no. 6 which are connected to each other with black lines.

The color red indicates the place of the first and second formants of the vowels of the speaker no. 7 which are connected to each other with red violet lines.

The color violet indicates the place of the first and second formants of the vowels of the speaker no. 8 which are connected to each other with violet lines.

The color yellow, indicates the place of the first and second formants of the vowels of the speaker no. 9 which are connected to each other with yellow lines.

The color blue indicates the place of the first and second formants of the vowels of the speaker no. 10 which are connected to each other with blue lines.

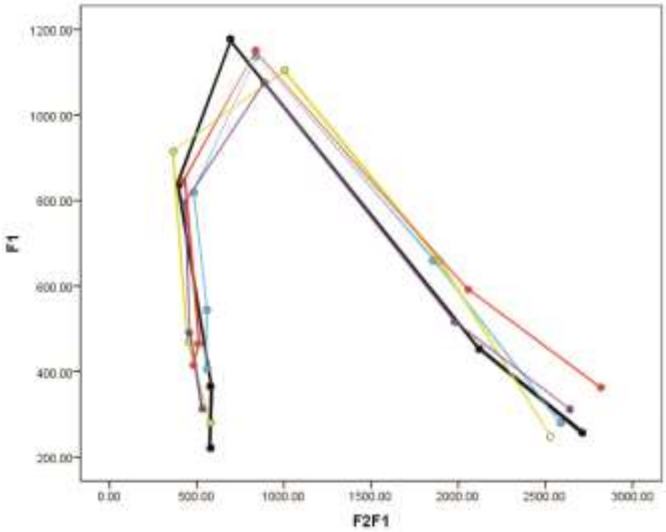

Figure 10: The exact vowel structure according to F2 and the distance between F1 & F2 for the speakers6-10 (female)

The color black indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 6 which are connected to each other with black lines.

The color red indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 7 which are connected to each other with red lines.

The color violet indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 8 which are connected to each other with violet lines.

The color yellow indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 9 which are connected to each other with yellow lines.

The color blue indicates the place of the first formant and the distance between the first and the second formant of the vowels of speaker no. 10 which are connected to each other with blue lines.

In order to make sure that the formant structure of the vowel of a single person in different contexts and different words has a similar pattern, another supplementary experiment in this study was conducted and based on this experiment the formants structure of the vowels of each female speaker in different words was extracted and its graph were compared to each other. In this experiment, the vowel formants were extracted from various sonic contexts and different places. The method of study and the word selection from amongst sets of data has also been random. The figure and table below, indicate the result of this experiment for speaker no. 6:

Vowel	The frequency of the formants of vowels pronounced by speaker 6 in different words		F1 Frequency (Hertz)	F2 Frequency (Hertz)	F3 Frequency (Hertz)	F2-F1 (Hertz)	F3-F2 (Hertz)
[ɩ]و اکه	The frequency of the	[αλιρεζΑ]	414	2369	3030	1955	661

	formants in vowel [ι]	[βΑζι]	390	2623	3233	2233	610
		[βιφΑ]	414	2623	3233	2209	610
		[μιζ]	439	2801	3512	2362	711
		[σινι]	490	2700	3207	2210	507
vowel[ε]	**The frequency of the formants in vowel [ε]**	[?εησΑσ]	684	2201	3008	1517	807
		[?εφτεμΑ?]	618	2166	2953	1548	787
		[?εστερεσ]	613	2166	3004	1553	838
		[κετΑβ]	592	2014	2852	1422	838
		[λεβΑσ]	490	2065	2928	1575	863
vowel [α]	**The frequency of the formants in vowel [α]**	[?αμμε]	1125	1709	2801	584	1092
		[?ανγοΣταρ]	1125	1709	2903	584	1194
		[?ανγυρ]	1125	1735	2750	610	1015
		[?αρυσ]	1100	1709	2750	609	1041
		[?ασβ]	1125	1709	2776	584	1067
vowel [υ]	**The frequency of the formants in vowel [υ]**	[βολανδγυ]	414	998	2750	584	1752
		[δυδ]	414	999	2877	585	1878
		[δυρ]	414	998	2725	584	1727
		[δυστ]	414	998	2700	584	1702
		[κΗαβυτΗαρ]	414	973	2547	559	1574
vowel [o]	**The frequency of the formants in vowel [O]**	[δοξτΗαρ]	439	1252	2826	813	1574
		[ξοδΑ]	490	1025	2725	535	1700
		[μοδιρ]	490	1024	2674	534	1650
		[οφτΗΑδ]	490	947	2953	457	2006
		[ροωΣαν]	490	1049	2700	559	1651
vowel [A]	**The frequency of the formants in vowel [A]**	[?ΑβΑδΑν]	795	1100	2598	305	1498
		[/ΑΓΑ]	769	1075	2649	306	1574
		[γΑη]	719	1151	2725	432	1574
		[κΗοφΑ]	719	1303	2674	584	1371
		[ζιβΑ]	719	1100	2750	381	1650

Table 11: The frequency of the formants of vowels pronounced by speaker 6 in different words

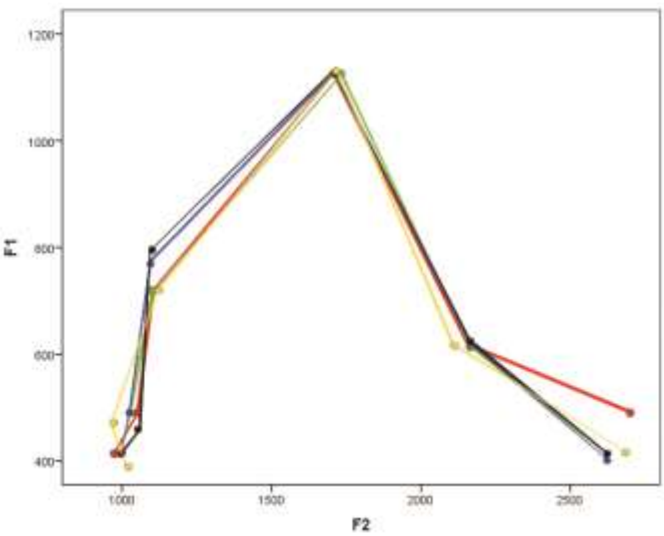

Figure 11: The graph related to the exact vowel structure according to F1, F2 for speaker 6 in 5 different sets of words.

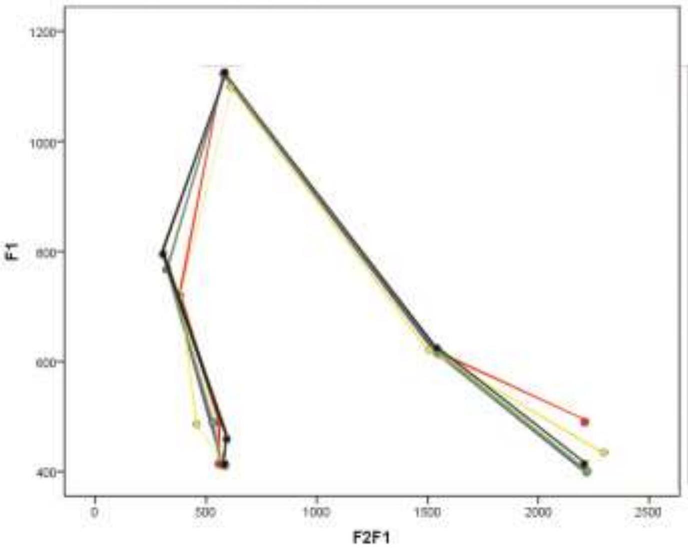

Figure 12: The exact vowel structure according to F2 and the distance between F1 and F2 for speaker 6 in 5 different sets of words.

As it is seen, although in the graphs mapped here the formant structures are also not in accordance with each other, the formant structure of this single speaker in various pronunciations have not changed so much. This test was conducted to all the speakers from 1 to 10 and gave the same results. Based on the results of these experiments, we can confirm the primary hypothesis of possible use of formant structure of vowels of a person for forensic speaker recognition.

4. Conclusion

The results of 24 separate experiments on this feature revealed that the hypothesis of considering this feature as a key for forensic speaker recognition is approved. This key is amongst the new and invaluable achievements gained in this thesis.

It is wise to mention that despite the limited numbers of the speakers whose sound data were studied, and for the vast volume of data being studied and experimented and the precise acoustics studies conducted from various aspects, and also the numbers of numerous keys which were used for recognition, we can say the findings of this study, enjoy with a plausible academic credit, but this doesn't mean at all that the keys and the features introduced in this thesis, are the only existing ones and the forensic speaker recognition studies can pave the way for the scholars and researchers toward extensive studies.

==

References

Aghagolzade, F. (2003), Firensic Linguistics, A new approach in applied linguistics, The first Iranian Conference of Linguistics Society, Tehran, Iran, Allameh Tabatabayi University Publications p: 215-225.

Bavelas , L. Gibson , (1994) J. ***Anomalous language in sexual assault trial judgments***, Discourse & Society, 5 (2): 189-206.

Harris, S. ***"Fragmented Narratives & Multiple Tellers: Witness & Defendant Accounts in Trials,"*** Discourse Studies, pp. 53-74.

Herbig, T. Gerl, F. & Minker, W. (2011) ***Evolution of an Adaptive Unsupervised Speech Controlled System***, IEEE Workshop on Evolving & Adaptive Intelligent Systems, EAIS-2011, Paris (France),

Hollein, H. (1990***), The Acoustics of Crime***, New York: Plenum Press. 190-191.

Kersta, L.G. (1962), **"Voiceprint identification"**. Nature, Vol. 196, No. 4861, pp.1253-1257.

Kinoshita, Y. (1998), "**Japanese *Forensic Phonetics: Noncontemporaneous Within-Speaker Variation In Natural & Readout Speech",*** in R. Mannell & J. Robert-Ribes (eds.) Proceedings of the 5th International Conference on Spoken Language Processing, Australian Speech Science & Technology Association, Canberra, pp. 145-148.

Langford, Ian. (2000). ***"Forensic semantics: the meaning of murder, manslaughter & homicide"***. Forensic Linguistics. 7.1, pp. 72-94.

Nolan, F. (1997), ***Speaker recognition & forensic phonetics***. In W. J. Hardcastle & J. Laver (Eds.), The H&book of Phonetic Sciences,744-767. Cambridge: Cambridge University Press.

Nolan, Francis, McDougall, Kirsty, Jong, Gea, & Hudson, Toby, (2006), ***A Forensic Phonetic Study of 'Dynamic' Sources of Variability in Speech: The DyViS Project,*** Department of Linguistics, University of Cambridge, United Kingdom.

Over, H. Lashbrook, W. Pedrev, C. Nicol, J. Nash, E.& Tosi, O. (1972), ***Experiment on voice identification.*** Journal of the Acoustical Society of America, 51 pp.2030-2043.

Rodman, R. D., (1998) ***Speaker recognition of disguised voices.*** In M. Demirekler, A.Saranli, H. Altincay & A. Paoloni (eds) Proceedings of the Consortium on peechTechnology Conference on Speaker Recognition by Man & Machine: directions for forensic application pp. 9–22. Ankara, Turkey: COST250 Publishing Arm.

Samareh, Y. ***Persian Phonetics: Phones and syllable Structure***, Jahad Daneshgahi Publication Center, Tehran, Iran. pp: 35-38.

Stevens, K.N. , Williams, C.E. , Carbonelli, J.R. & Woods, B. ,(1968), **"Speaker authentication & identification: a comparison of spectrographic & auditory presentations of speech material",** Journal of the AcousticalSociety of America, 43 pp.1596-1607.

Tiersma, M. P. (1999), **Legal Language** Chicago: University of Chicago Press.

==

Mojtaba Namvar Fargi
mojtabanamvar@gmail.com

Shahla Sharifi
sh-sharifi@um.ac.ir

Mohammad Reza Pahlavan-Nezhad
Rezapahlavan5@yahoo.com

Azam Estaji
Estajiz@yahoo.com

Mehi Meshkat Aldini

Linguistics Department
Ferdowsi University of Mashhad
PO box 91779-48974
Mashhad
Iran