MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Influence of classroom acoustics on student presentations

Subramaniam, Nithya[1]
Indian Institute of Technology, Madras
Chennai – 600 036

Ramamurthy K.[2]
Indian Institute of Technology, Madras
Chennai – 600 036

ABSTRACT

Teachers are better speakers when compared to students with oracy built into their professional experience over a period of time. Students' perception and judgement of teachers' speech intelligibility has been reported abundantly in literature. However, teachers' perception of students' speech intelligibility has been rarely studied in the context of classroom acoustics. Graduate students deliver presentations and talks regularly in classrooms and are assessed and also graded based on their communication skills and delivery. Vowel space area (VSA) is an objective metric to evaluate the talker speech intelligibility and articulation. From recorded presentations, VSA has been evaluated from the first two formant frequencies, F_1 and F_2 measured from speech waveforms of words containing the corner vowels. VSA analysis was done in two different classroom acoustical conditions. Acoustical conditions were characterized for the two classrooms by measuring octave band reverberation times and background noise levels. This was used to evaluate the room average useful-to-detrimental energy ratio (U_{50}), an objective metric to evaluate intelligibility in rooms for speech. Influence of classroom acoustical characteristics on students' speech intelligibility was studied by comparing VSA's for all student talkers and the classroom average U_{50} metric. The results indicate that teachers' judgment of students' speech and delivery could very well be influenced by classroom acoustical conditions.

Keywords: Vowel space area, Useful-to-detrimental ratio, Intelligibility
I-INCE Classification of Subject Number: 63

1. INTRODUCTION

Acoustical comfort in general and quality of speech production and transmission from students to teachers in particular are important in determining the impressions students make on their teacher. At similar talker-listener distances, but in varying room environments, for the same talker, the listeners' perception may be modified due to room effects or due to talkers' modified speech.

[1]nithyaalways@gmail.com
[2]vivek@iitm.ac.in

The type of speaker and environment are the main factors of influence [1]. There are likely to be individual differences in the strategies used by talkers to clarify their speech in different speaking conditions, as well as the degree of communication success they achieve [1, 2]. While it is widely acknowledged that speakers greatly vary in the intelligibility through speech modifications that happen due to changes done to speech production and delivery [3,4], very few studies have looked at how talkers' intelligibility differ when subject to different acoustical environments. Moreover, not many studies have been carried during actual real classroom sessions using live speech recording samples for study.

Bradlow et al. [5], in their work have established the correlation between speech intelligibility scores and global and fine-grained acoustical characteristics such as fundamental frequency range, vowel space measures etc. Vowel analysis from segmental study of speech is useful for predicting the speech intelligibility of talkers. Research supports the belief that intelligibility is linked to the properties of the first and second vowel formant frequencies F1-F2 [6, 7]. Although it is still unclear as to which properties of the vowel space correlate the best with intelligibility, overall size of the vowel space is found to be a reliable predictor of intelligibility [8].

From the room acoustics perspective, useful-to-detrimental ratio (U_{50}) gives the relationship between beneficial sound energy from the talker and the reverberant sound including the noise. The U_{50} measure combines the detrimental effects of late-arriving speech and ambient noise relative to the useful direct and early reflected speech sounds. A minimum U_{50} of +1 dB is recommended for satisfactory intelligibility [9].

In order to evaluate whether student-talkers would improve their intelligibility when they presented seminars in rooms of lower U_{50} value, all talkers were evaluated when speaking in two classrooms having different physical and acoustical characteristics. Relationships between room acoustic and acoustic phonetic correlates of intelligibility if explored could have practical implications for developing speech enhancement technologies to improve listener perceptions and also for judiciously modifying the acoustical environment of classrooms to enhance its acoustical quality.

2. METHODOLOGY

2.1 Classrooms

The two classrooms chosen were different in terms of volume, interior surface properties and furnishings. Table (1) depicts the physical characteristics of the classrooms studied.

Table (1) – Physical characteristics of classrooms

Quantity	Classroom 1	Classroom 2
Room Volume (m^3)	396.4	494.6
Maximum Source-receiver distance	7	10
Room width (m)	9.1	9.1
Length (m)	9.2	12.1
Height (m)	3.6	3.6
Surface area (m^2)	372.8	449
Volume/Surface area (m)	1.05	1.1
Seating capacity (Num)	120	156

In order to characterise the acoustical quality of the classrooms used by students for presentations, 1/1 octave band measurements of reverberation times and background noise levels were carried out in two classrooms. Reverberation times were measured using Balloon pop impulse source method. The octave band frequencies in the range of 125 Hz – 8000 Hz were excited and the reverberation time (T_{20}) was measured using the integrating sound level meter Norsonic Type 118. Figure (1) shows the typical receiver locations selected for measuring the room acoustical parameters.

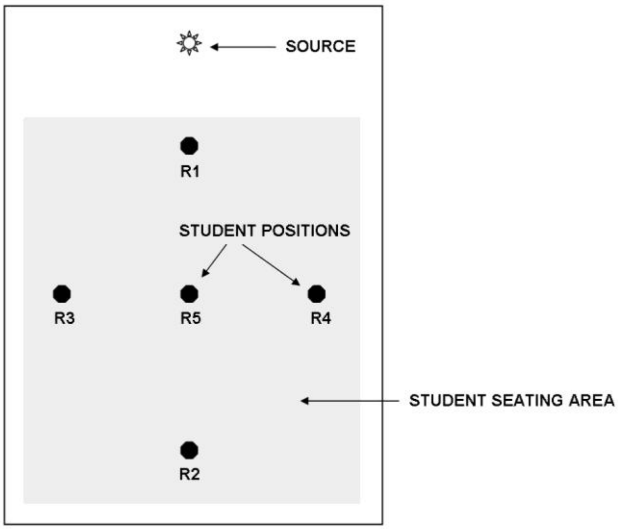

Figure 1 – Typical measurement locations in the classroom

The fenestration was kept open and the mechanical ventilation systems were operating as per normal in-use classroom conditions throughout the entire experiment. From the measured impulse responses in both classrooms 1 and 2, useful-to-detrimental ratios (U_{50}) were derived. The following equation by Bradley [10] was used to arrive at 1 KHz - U_{50} values at typical receiver locations (figure 1) in the classroom. The 1 KHz values were averaged spatially to get the classroom average U_{50}. U_{50} is given by equation (1) as;

$$U_{50} = \left[\frac{10\log\left(\frac{1+r_h^2}{r^2}\right) - e^{\left(\frac{-0.69}{RT}\right)}}{e^{\left(\frac{-0.69}{RT}\right)} + 10^{\frac{SNR}{10}}} \right] \qquad - \qquad \text{(Equation 1)}$$

where, 'r_h' is the reverberation distance where the energy densities of the reflected and direct sounds become equal; 'r' is the source-receiver distance; 'RT' is the reverberation time; SNR is the signal-to-noise ratio at 'r'.

2.2 Student talkers and presentation conditions

10 student talkers (T1 to T10) were randomly selected with equal number of males and females in the age group of 22 to 30 years. The talkers assumed the typical lecturing position on the dais in the classrooms and delivered their talks. Microphone was placed at 5 cms from their mouth and their speech was recorded for the entire duration of their talk. The microphone output was connected to the PC and the

recordings were made directly using the Praat software interface [11]. The talkers were instructed to speak naturally. It was made sure that no talker presented twice on the same day. The experiments were conducted in two classrooms on separate days for each talker.

2.3 Vowel space measurements

Speech intelligibility improvements are always accompanied by a wide range of acoustic changes. Previous studies suggest that larger vowel space areas are associated with increased intelligibility [5]. Four-point vowel spaces for all the ten talkers in the study are evaluated for the two classroom acoustical conditions in which speech was elicited.

The waveforms for words containing the corner vowels [a], [ɑ], [i] and [u] that make up the quadrilateral were segmented from running speech samples recorded during presentations in classrooms 1 and 2, using Praat software. With Praat, the vowel segment was selected and saved for further analysis. Praat provides the parallel display of the speech waveform, spectrogram and the short spectral cross-section. Formant frequencies F1 and F2 were measured manually, by placing the mouse cursor at the steady state portion of the vowel segment and values were recorded. Five tokens for each vowel case were measured for each of the ten student-talkers and this works out to 200 manual measurements totally. The mean F1 or F2 value of the five tokens per vowel is taken as the formant value in Hertz for the particular vowel.

The formant frequency measures in Hertz were converted into the psycho-acoustical Bark scale using the equation given by Traunmuller [12],

$$F\,(\text{Barks}) = \left[\frac{26.81}{\left(1+\left(\frac{1960}{F(\text{Hz})}\right)\right)}\right] - 0.53 \qquad - \qquad \text{(Equation 2)}$$

where F (Hz) is the value of the formant frequencies in Hertz and F (Barks) is the transformed Bark value. From these mean Bark scale formant measures, the F1, F2 coordinates for each vowel was fixed and the area of the polygon formed from the F1,F2 co-ordinates of each vowel was derived. The area of the polygon is the vowel space area (VSA). VSA's are compared across talkers from speech samples in two different classroom acoustical conditions. The variation in vowel space characteristics for the same talker across classrooms is also studied.

3. RESULTS AND DISCUSSION

3.1 Room acoustical characteristics

The 1 KHz room average U_{50} values evaluated from Equation (1) are – 1.9 dB and – 2.5 dB for classrooms 1 and 2 respectively. Although there is a difference of less than 1 dB in the 1 KHz - U_{50} values between classrooms 1 and 2, there is a perceptible difference in their acoustical qualities with classroom 2 being smaller and less reverberant and noisy in the low frequency bands as seen in Figures (2) and (3).

The mid and high frequency reverberation times only differ by 0.2 seconds between the rooms. However, in the low frequency bands, there is a very significant difference in RT and even more so in the case of background noise levels with a difference greater than 5 dBA between the rooms. The L_{Aeq} levels in classrooms 1 and 2 were 60.8 dBA and 65.2 dBA respectively.

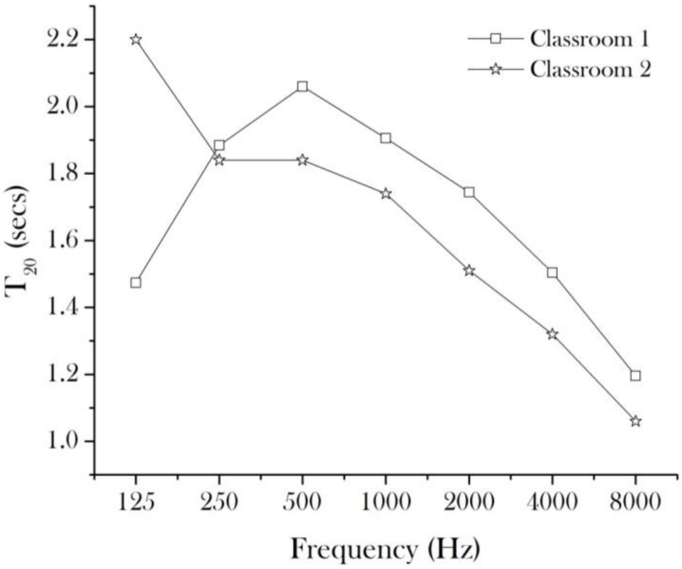

Figure 2 – Octave band reverberation times for classrooms 1 and 2

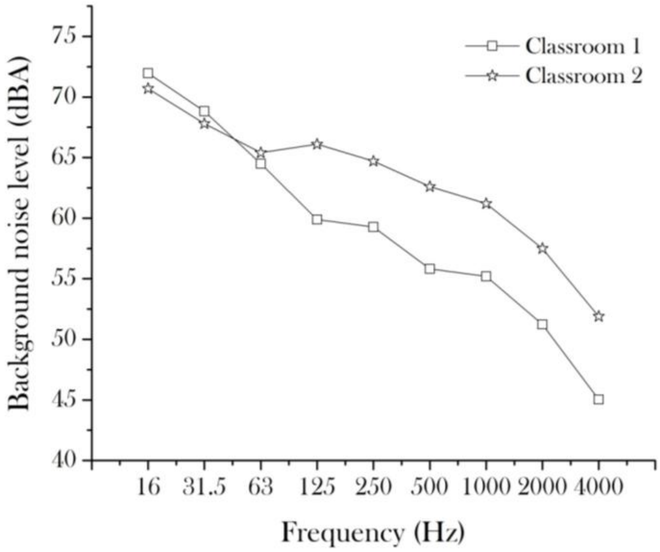

Figure 3 – Background noise levels in classrooms 1 and 2

Classroom 1 is more reverberant at mid and high frequencies when compared to classroom 2, whereas classroom 2 is noisier of the two. The predicted intelligibility in classroom 2 is comparatively lower as suggested by the U_{50} metric. The present investigation compares the gross acoustic-phonetic changes in terms of vowel production that occur when speech is delivered in rooms having different acoustical characteristics.

3.2 Vowel space characteristics of student-talkers

The vowel space areas range between 6.5 to 14.1 bark units in classroom 1 and between 8.3 and 13.2 bark units in classroom 2. Figures (4) and (5) depict the vowel space areas for all talkers in classroom 1 and 2 respectively.

In classroom 1, the F1 formant frequency across talkers varies from 3 to 9.5 Barks and F2 formant frequency varies from 7.5 to 15 Barks. In classroom 2, the F1 values range between 3 and 9 Barks and the F2 values range between 7 and 15 Barks..

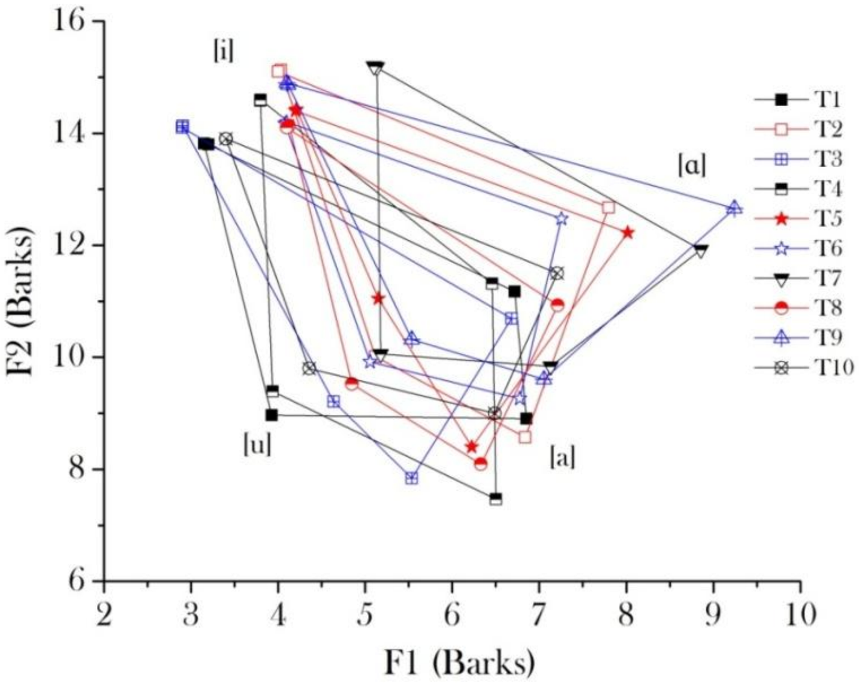

Figure 4 – Vowel space areas (bark units) for all student talkers in Classroom 1

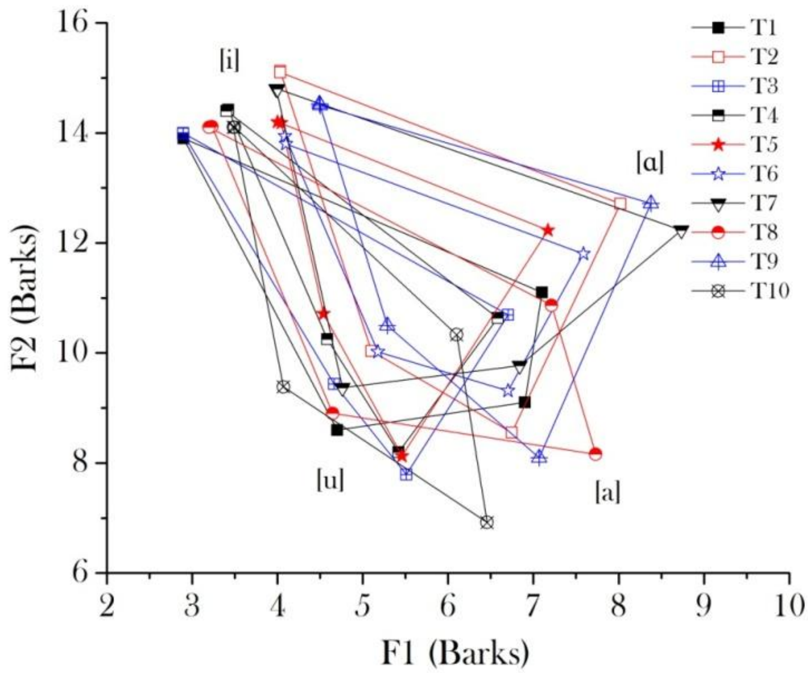

Figure 5 – Vowel space areas (Bark units) for all student talkers in Classroom 2

Table (2) shows the range of F1 and F2 values for each vowel type across all talkers. Although the overall range of formants across classrooms is similar, they vary across talkers considerably and in between rooms as discussed below.

Table (2) – Range of F1 and F2 values (barks) across all talkers

Class room	[i]		[a]		[u]		[ɑ]	
	F1	**F2**	**F1**	**F2**	**F1**	**F2**	**F1**	**F2**
1	3 - 5	13 – 15	5.5 – 7	7.5 – 10	4 – 5.5	9 – 11	6.5 – 9	10.5 – 13
2	3 – 4.5	13.5 – 14.5	5.5 – 7.5	6.5 - 10	4.5 - 5.5	8 - 10.5	6 - 9	10 - 13

VSA is found to increase in classroom 2 for most talkers as can be seen in figure (6).This could possibly be due to the modifications made to their articulation and speech in order to enhance their intelligibility in a comparatively degraded listening environment. Since larger VSA's are associated with higher intelligibility, it can be seen that there is a trend for majority of talkers to enhance their intelligibility by expanding their vowel working areas when moving from a classroom with higher U_{50} to one with lower U_{50}.

Talkers T1, T3 T9 and T10 show marginal increase in vowel space areas, whereas talkers T5 and T6 show a moderate increase. Talker T4 shows a significant increase in vowel space area in classroom 2 with higher lower U_{50}. On the other hand, there are few talkers who exhibit a reduction in their VSA's with talker T2 showing a small reduction in VSA and talkers T7 and T8 showing substantial reduction in VSA.

Although the difference in U_{50} may not correspond to perceptible changes in listener perception in a given environment, it suggests perceptible changes in talker intelligibility in terms of vowel space measures as indicated by Figure (6).

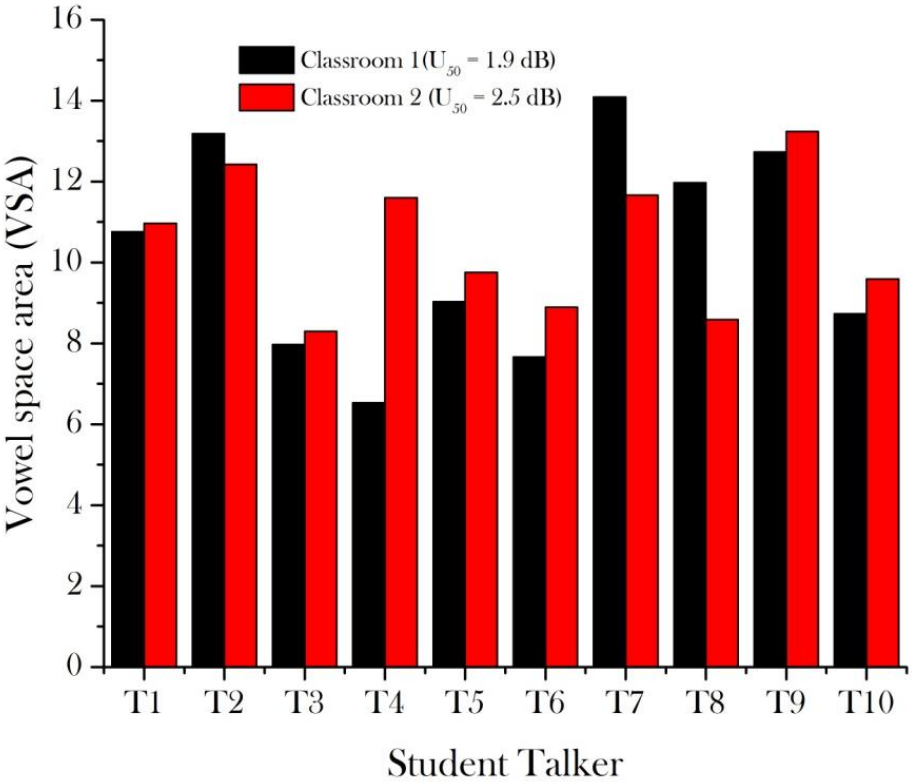

Figure 6 – Vowel space areas in classrooms 1 and 2 for all talkers

Table (3) and (4) gives the difference in values of F1 and F2 between classrooms respectively for all corner vowels for all the talkers. Comparatively, changes in F2 are much less across talkers. From the data, it can be gauged that expansion in the F1 region is more prominent when exposed to different room acoustical conditions.

Table (3) – Comparison of difference in F1 for all talkers between classrooms 1 and 2

Talker	**Difference in Formant frequency F1 (barks) between Classrooms 1 and 2**			
	[i]	**[u]**	**[a]**	**[ɑ]**
1	0.3	0.8	0	0.4
2	0	0	0.1	0.2
3	0	0.1	0	0
4	0.9	0.2	1.4	0
5	0.2	0.7	0.7	0.8
6	0	0.1	0.1	0.3
7	1.1	0.4	0.3	0.2
8	0.9	0.2	1.4	0
9	0.4	0.2	2.1	4.3
10	0.1	0.3	0	1.1

Table (4) – Comparison of difference in F2 for all talkers between classrooms 1 and 2

Talker	**Difference in Formant frequency F2 (barks) between Classrooms 1 and 2**			
	[i]	**[u]**	**[a]**	**[ɑ]**
1	0.1	0.3	0.2	0
2	0	0	0	0
3	0.2	0.2	0.1	0
4	0	0.6	0	0.1
5	0.3	0.3	0.3	0
6	0.3	0.1	0	0.7
7	0.4	0.8	0.1	0.3
8	0	0.6	0	0.1
9	0.4	0.2	1.6	0
10	0.2	0.5	3.9	1.2

With respect to articulatory gestures and talkers' tongue positions when speaking in different classroom environments, the F1 and F2 values are found to vary within talkers. From tables (3) and (4) it can be seen that the differences in F1 are more compared to differences in F2 for a given talker. However, the patterns of formant movements are different across vowels. The mean difference in F1 across talkers is more for vowel [ɑ] and [a] when compared to vowels [i] and [u]. In the case of F2, the mean difference is greater for vowel [u]. The mean difference is more for vowel [u] in the case of both formants F1 and F2. The largest difference is noted in the case of F1 for vowel [ɑ].

The frequency of the first formant increases as we open our mouth wider and lower the tongue. The frequency of F2 increases as we advance our tongue. This influences the vowel space working areas in two different classroom environments and two groups of student talkers are identified, one group with 4 out of 10 talkers who clearly showed no difference in VSA or deterioration in intelligibility in the classroom with lower $U_{50,}$ termed as talker group A and another group of 6 who showed considerable speech intelligibility improvements in terms of vowel space areas with expansions in F1 region termed as talker group B.

3.3 Talker group A

From Figure (7) showings VSA's of talker group A below, it can be seen that Talkers 1,2 3 and 6 exhibit consistencies when speaking in different room acoustical conditions, without changing their formant characteristics. This indicates that these talkers are not affected by the environment in which they speak in and do not feel the need adapt to the constraints offered by the environment. Among talkers in group A, talkers 1 and 2 have comparatively higher vowel space areas than talkers 3 and 6 who exhibit compact vowel spaces.

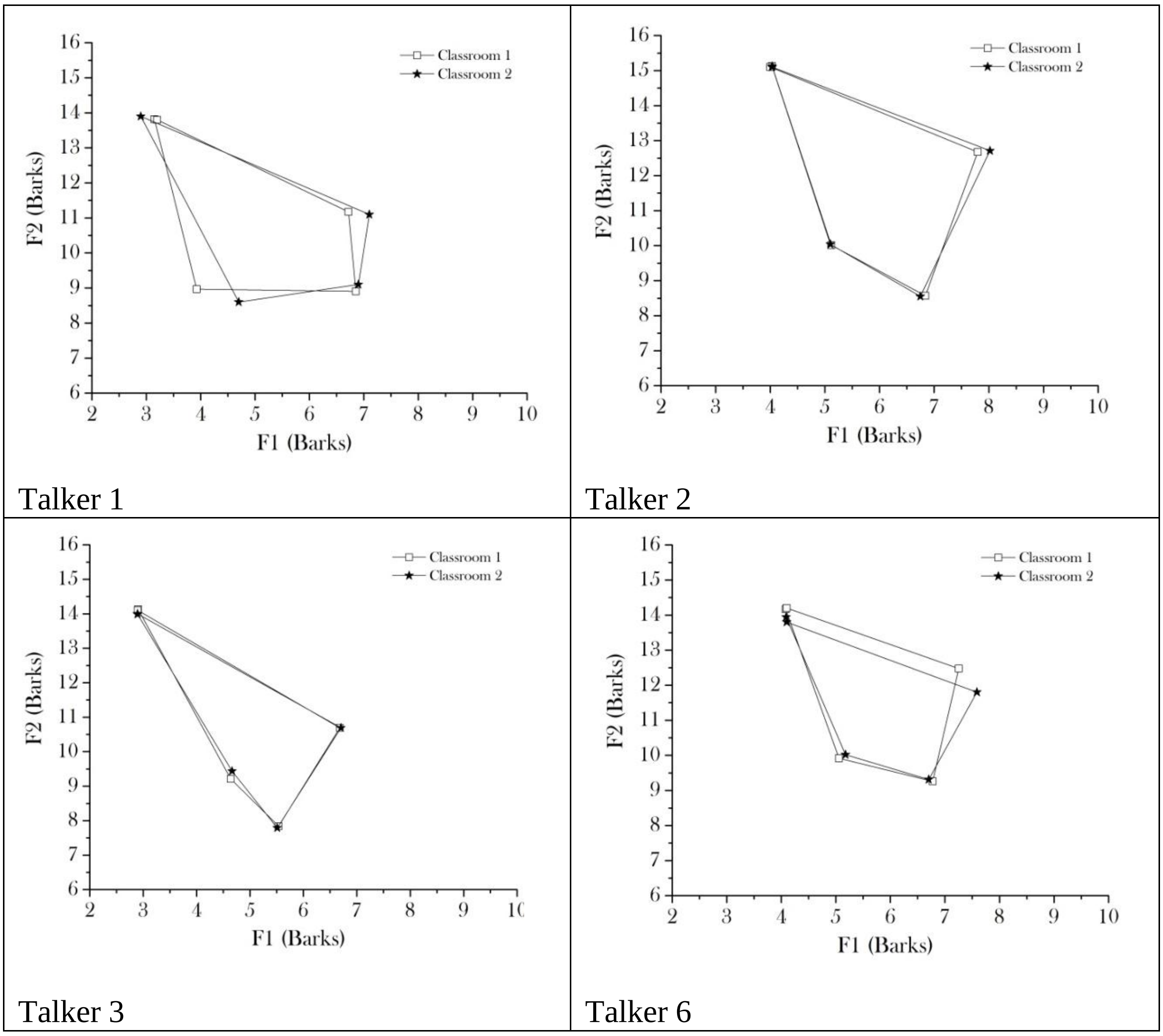

Figure 7 – Vowel space areas for student talker group A

3.4 Talker group B

Although vowel space areas do not differ much in certain talkers such as T5, T9 and T10 as indicated in figure (8) showing VSA's for group B, the formant frequencies are modified considerably as can be seen in their individual vowel spaces. Except for talkers 9 and 10, the remaining talkers in group B exhibit expansions or reductions in the F1 dimension. In articulatory terms, this suggests a wider or narrower opening of the mouth. This considerably can reduce or improve the clarity or intelligibility of speech produced.

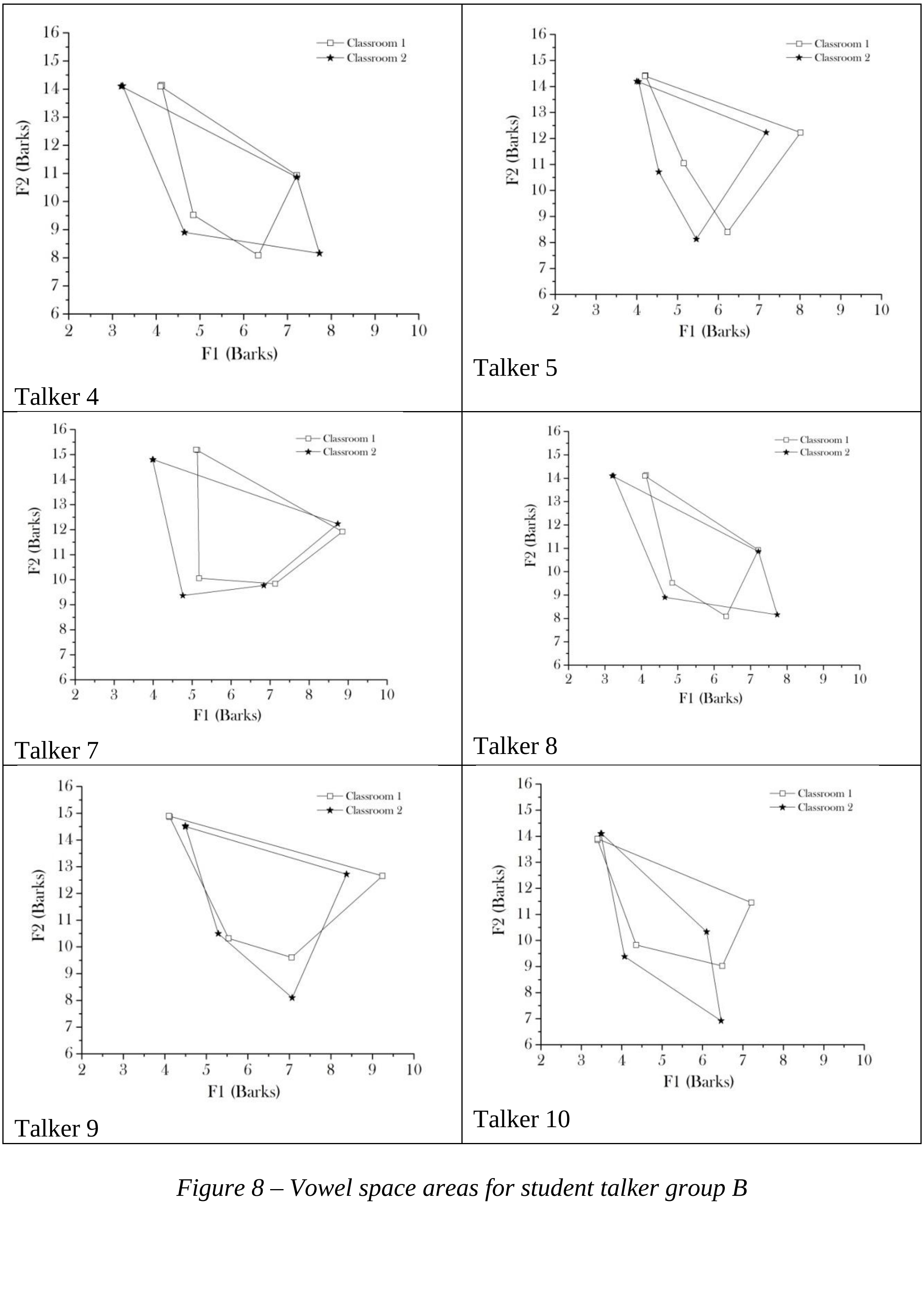

Figure 8 – Vowel space areas for student talker group B

4. CONCLUSIONS

Although majority of the talkers show an increasing vowel space expansion trend in a classroom with lower U_{50}, vowel space expansions were not observed uniformly within the group of talkers. The results indicate the multi-dimensional nature of speech and considering a wide range of acoustic-phonetic characteristics that can vary during speech production and delivery, the vowel space expansion for some of the talkers in one room compared to the other was interesting to observe. Moreover talkers with more or less equal vowel spaces showed differences in formant frequencies expansions. The Euclidean distance between the vowel formants for both room conditions tested could give a better perspective of the intelligibility enhancements made by individual talkers.

This suggests that strategies employed by some student talkers to improve their intelligibility by making specific changes to their vowel production when talking in different room acoustical conditions may have the desired effect on their speech performance and delivery. Thus, room acoustical characteristics have the potential to influence student-talkers' speech and consequently the impression created on their teacher.

5. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support from all student volunteers who participated in the experiment.

6. REFERENCES

1. Hazan, Valerie, and Rachel Baker. "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions." *The Journal of the Acoustical Society of America* 130.4 (2011): 2139-2152.
2. Krause, Jean C., and Louis D. Braida. "Acoustic properties of naturally produced clear speech at normal speaking rates." *The Journal of the Acoustical Society of America* 115.1 (2004): 362-378.
3. Ferguson, Sarah Hargus, and Diane Kewley-Port. "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners." *The Journal of the Acoustical Society of America* 112.1 (2002): 259-271.
4. Bradlow, Ann R., Nina Kraus, and Erin Hayes. "Speaking clearly for children with learning disabilities." *Journal of Speech, Language, and Hearing Research* (2003).
5. Bradlow, Ann R., Gina M. Torretta, and David B. Pisoni. "Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics." *Speech communication* 20.3-4 (1996): 255-272.
6. Bond, Zinny S., and Thomas J. Moore. "A note on the acoustic-phonetic characteristics of inadvertently clear speech." *Speech communication* 14.4 (1994): 325-337.
7. Hazan, Valerie, and Duncan Markham. "Acoustic-phonetic correlates of talker intelligibility for adults and children." *The Journal of the Acoustical Society of America* 116.5 (2004): 3108-3118.
8. Amano-Kusumoto, Akiko, and John-Paul Hosom. "A review of research on speech intelligibility and correlations with acoustic features." *Center for Spoken*

Language Understanding, Oregon Health and Science University (Technical Report CSLU-011-001) (2011).

9. Bradley, John S. "Speech intelligibility studies in classrooms." *The Journal of the Acoustical Society of America* 80.3 (1986): 846-854.
10. Bistafa, Sylvio R., and John S. Bradley. "Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics." *The Journal of the Acoustical Society of America* 107.2 (2000): 861-875.
11. Boersma, Paul, and David Weenink. "Praat: Doing phonetics by computer, version 4.0. 26." *Retrieved September* 24 (2005)
12. Traunmüller, Hartmut. "Analytical expressions for the tonotopic sensory scale." *The Journal of the Acoustical Society of America* 88.1 (1990): 97-100.