Crowd Noise and Vocal power level in Large College Canteens in China

Chao Wang[c], Xuejiao Kong[b, a], Sheng Yao[a], Jian Kang[c, d], *and* Jingyu Yuan[a, *]

a: School of Architecture and Art Design, Hebei University of Technology, 300401 Tianjin, China

b: Department of Digital Arts, Tianjin Electronic Information College, 300350 Tianjin, China

c: School of Architecture, Tianjin University, 300072 Tianjin, China

d: UCL Institute for Environmental Design and Engineering, The Bartlett, University College London (UCL),Central House, 14 Upper Woburn Place, London WC1H 0NN, UK;

**: corresponding author. jyyuan@hebut.edu.cn; yaosheng@hebut.edu.cn; Tel.: +86-022-6020-1059; Xiping Road No.5340, Beichen District, Tianjin, 300401, China.*

Abstract:

Both the noise and the vocal power level of the crowd are parameters for evaluating the acoustic environment. The former is from the perspective of the environment, and the latter is from the perspective of the user's needs. This paper aims to explore the crowd noise and vocal power level of diners in large college canteens in China. Measurements were conducted in two typical Chinese college canteens. Videos were also recorded to analyse the diners' behaviour. The results showed that the noise in canteens varied from 61 dBA to 73 dBA during the meal. It was noted that although the noise level had a strong correlation with the number of occupants, the relationship between this two was not simply the superposition of equal-intensity sound sources. The speaking ratio and the Lombard effect played a significant role during the meal. The average speaking ratio was 0.12, which was far less than the practical value, 1/3. A prediction model for vocal power level under crowd noise was introduced into Chinese college canteens, which considered direct and reverberant sound energy using the parameters of room information, the location of the diners, and the speaking ratio. Based on this model, the vocal power level was found to vary from 59 dBA to 97 dBA. According to Pearson's evaluation criterion, 17.1% of the data fell between "Raised" and "Loud", and 71.8% of the data fell between "Loud" and "Shout". It indicated that the crowd noise in Chinese college canteens had a significant impact on communication and some acoustic treatment was necessary.

Keywords: Crowd noise, Vocal power level, Lombard effect, Speaking ratio, Chinese college canteen, Mandarin

1 Introduction

This paper aims to explore the crowd noise and vocal power level of diners in large college canteens in China. Chinese college canteen usually serves a large number of students during the eating time [1], which is more like a large and daily-used banquet hall with hard walls and ceiling. Studies on crowd noise in large college canteens are useful for acoustic design and noise control, which will enhance the acoustic environment of dining spaces and benefit the students and staffs in colleges.

The noise in the dining space was usually very high [2, 3], and it became more serious in Chinese college canteens. This continuous high-noise-level environment made it difficult to communicate and caused low comfort and satisfaction [4, 5] of the diners. Several measurements [1, 6, 7] showed that Chinese college canteens usually had long reverberation time and the noise level varied from 70 to 80dBA, which lasted nearly two hours per meal. The noise in the eating establishments can be divided into two parts [8]. One is the "physical" background noise (including the noises from kitchen equipment, HVAC equipment, etc.), and the other is the crowd noise, which domains in most cases and varies with the number of occupants.

Crowd noise is a particular type of noise that its source and receiver are both the people [9]. People will unconsciously enhance their voice in a noisy environment[10, 11], which not only reduces the comfort of the diners but also increases the severity of the noise. This phenomenon was called the "Lombard effect" [12, 13, 14]. As the background noise rises gradually from silence, the phenomenon of the Lombard effect can be divided into three stages. First, the Lombard effect is not apparent when the noise is too low. And then, the output level of an individual's voice starts to increase at a certain point [15] and then continuously increases. Finally, it no longer increases due to reaching the individual's vocal limit [16]. The Lombard slope varied from 0.2~0.7 dB/dB in different situations [17, 16]. When a new person enters a space, he will be affected by noise and begins to raise his voice to maintain communication [18]. The noise generated by this person also affects everyone else in the space, resulting in a very short iterative process until it converges again. This iterative process frequently occurs during the entire dining period.

Vocal power level was an important parameter to evaluate the vocal effort of an individual [19] and also represented the difficulty of communication. Based on Pearson's evaluation criterion [20], the vocal power level was classified into normal, raised, loud and shout. It is also important for the noise prediction because most of the crowd noise prediction in the eating establishments can be described by two steps: determining an individual's vocal effort and calculating the noise accumulation of all the diners based on different assumptions [21]. However, the vocal power level is influenced by many factors, which made it difficult to be predicted accurately. For example, the composition ratio of gender and age would affect the intensity and frequency distribution of an individual's vocal effort [22]. Situational Context and whether or not serving alcohol could also influence it [23]. Group size is also an important factor, which effect could be divided into two aspects as follows. It is reasonable that only one speaker existed in a group at the same time, and theoretically, the speaker's sound power can be evenly divided by everyone in the group. So, the more massive the group size is, the less vocal energy each person could average. Nonetheless, a larger group size also means a more considerable average distance between the speaker and listener, which will cause the speaker to increase the voice [24]. The average group size varied in different situations and was 3 to 4 in most eating establishments [23, 25].

Deriving from the on-site noise in a space is a particular method to obtain the vocal power level of the crowds, which is suitable for some situations where exists high noise or the microphone is not appropriate to approach the human mouth. It also avoids the distortion of the test environment in the laboratory [26]. Several models have been proposed to predict the vocal power level in eating establishments. Hodgson [27] investigated the relationship between the noise level and the number of occupants in ten eating establishments through questionnaires and measurements. And then, optimisation techniques were used to find the best estimate of unknown prediction parameters, using the measured noise level as the prediction target. Rindel [28] explored the vocal power level and speech signal-to-noise ratio in restaurants, and the concept of “acoustic capacity” was proposed based on reverberation sound energy superposition and computer simulation technology. However, all the models above were based on the assumption of the diffuse sound field, and only the reverberant energy was considered. There were few considerations for the direct sound in the prediction of vocal power level. Njis [16] explored the role of direct sound in crowd noise through the calculation in a virtual space, which showed that the the direct sound had a great impact on the result in the non-reverberant field. Tang [29] derived a method for calculating the direct sound accumulation of crowd to predict the noise level in an enclosed space. This implied a method of predicting vocal power level from the measured noise that takes into account both the direct sound energy and reverberant sound energy, which will be explored in this article.

The speaking ratio refers to the proportion of the number of speakers to the total number of occupants. This parameter was first used in the prediction model of crowd noise [27]. It is a critical parameter in the relationship between the noise level and the speaker's vocal effort [30]. In earlier models for crowd noise, the speaking ratio was used as a default value, 1/3 [27, 28]. However, noise can affect people's willingness to speak, so as the noise changes during a meal in the college canteen, the speaking ratio will also have a temporal difference. People also behave differently in different scenes, which will make the speaking ratio in the college canteen different from that in the restaurant or other spaces. Moreover, the speaking ratio may also vary in different cultural backgrounds. The long-term noise spectrum in 12 languages [31] found that the noise of different languages was similar, but the anti-interference ability was not the same. With the same output power, Mandarin was considered to have lower speech intelligibility [32] than English in noisy conditions, and Chinese may have higher noise sensitivity [33], which may also result in the difference in responses to noise. Therefore, the value of speaking ratio in Chinese college canteens and how it changes with factors such as noise level, gender and group size has not been determined by empirical investigations.

In this paper, a prediction model for vocal power level based on the on-site crowd noise was introduced into Chinese college canteens, which considers both the direct and reverberation sound. And then, two on-site measurements were conducted to determine the parameters, such as noise level, the number of diners, and the speaking ratio (by video records). Based on this prediction model, Lombard slope and vocal power level were obtained.

2 Method

2.1 Model Establishment

The vocal power level can be obtained from the noise level in the college canteens through the following model. Firstly, the crowd noise was obtained by subtracting background noise from total noise, as the total noise level

$$SPL_{total} = SPL_{ph} \dotplus SPL_{cr} \tag{Equation 1}$$

where, SPL_{ph} is "physical" noise level, which is often considered as a constant, as explained in the Introduction in this paper; SPL_{cr} is the crowd noise level, which can be affected by many factors, and its prediction is the key to most research problems; $\dotplus$ is the adding of energy. So the crowd noise

$$SPL_{cr} = SPL_{total} \dot{-} SPL_{ph} \tag{Equation 2}$$

where, $\dot{-}$ is the subtracting of energy.

Assuming that everyone has the same average vocal power level (in the entire dining process, including talking time and silent time), $L_{W,each}$, then the direct noise level at receiver j from all the speakers

$$SPL_{cr,j,dir} = L_{W,each} + 10\log_{10}\left[\sum_{i=1}^{n_{total}}\left(\frac{Q}{4\pi r_{ij}^2}\right)\right] \tag{Equation 3}$$

where, Q is the directivity factor, which is usually assumed to be 1 [25, 29], that is, the crowd is considered as a non-directional sound source; r_{ij} is the distance from the i-th sound source to the receiving point j; n_{total} is the total number of occupants. The reverberant noise level at receiver j from all the speakers

$$SPL_{cr,j,rev} = L_{W,each} + 10\log_{10}\left[\frac{4*n_{total}}{R}\right] \tag{Equation 4}$$

where R is the room constant. So, the noise level at receiver j from all the speakers

$$SPL_{cr,j} = SPL_{cr,j,dir} \dotplus SPL_{cr,j,rev}$$

$$= L_{W,each} + 10\log_{10}\left[\frac{Q}{4\pi}\sum_{i=1}^{n_{total}}\left(\frac{1}{r_{ij}^2}\right) + \frac{4*n_{total}}{R}\right] \qquad \text{Equation 5}$$

A similar relationship was also described by Tang [29] in his model for noise prediction. If the $SPL_{cr,j}$ was measured at point j, the average vocal power level of each diner

$$L_{W,each} = SPL_{cr,j} - 10\log_{10}\left[\frac{Q}{4\pi}\sum_{i=1}^{n_{total}}\left(\frac{1}{r_{ij}^2}\right) + \frac{4*n_{total}}{R}\right] \qquad \text{Equation 6}$$

If people are assumed to be evenly distributed, the distance between the diners

$$D_0 = \sqrt{\frac{S}{n_{total}}} \qquad \text{Equation 7}$$

where S is the area of the canteen. Then, the coordinates of each diner can be determined. In reality, there is a minimum distance between the diners, so the speaker closer than 0.5 m from the receiver point was eliminated to avoid the occurrence of the extreme value of sound power.

However, $L_{W,each}$ is not the vocal effort of the diners because they are not always talking during mealtime. So, the vocal power level when speaking

$$L_{W,\text{speaking}} = L_{W,each} \div \alpha \qquad \text{Equation 8}$$

where, $\div$ represents energy division. α is the speaking ratio.

How to obtain the speaking ratio is another difficulty. There are two ways: one is dividing the number of diners speaking by the total number of diners, and the other is dividing one person's speaking time by the total mealtime. Assuming that people speak "evenly" during a meal, then the two ratios are approximately equal, i.e.

$$\frac{n_{\text{speaking}}}{n_{total}} = \alpha_{\text{speaking}} \approx \alpha_{\text{speaking},eating} = \frac{\sum_{i=1}^{n_{total}} t_i}{\sum_{i=1}^{n_{total}} t_{total}} \qquad \text{Equation 9}$$

where, α_{speaking} is the ratio of people who are talking at a given moment; $\alpha_{\text{speaking},eating}$ is the ratio of the speaking time to the mealtime; n_{speaking} is the number of occupants who are speaking; t_i is the sum of the speaking time of a person during mealtime; t_{total} is the mealtime of a diner. It is very difficult to obtain α_{speaking} directly, so it was counted instead of $\alpha_{\text{speaking},eating}$ in this article. Considering everyone's mealtime is different, the noise level will also change during this period. However, if sufficient samples are recorded, a reasonable approximation of the speaking ratio can still be obtained.

When the noise level at each point has been measured, the crowd noise can be calculated by Equation 2. And then, the vocal power level can be deduced by Equation 6. Finally, the Lombard slope can also be obtained based on the relationship between the vocal power level and the noise level. Since the noise level used in this model is measured, there is no need to consider the iterative process of interaction between diners in the calculation process.

Compared with Rindel [28] and Hodgson's [27] models for predicting vocal effort, this model considers the influence of direct sound, which increases the complexity of the calculation. And two more results are predictable: (1) The addition of the direct sound will lead to a lower vocal effort result when the vocal effort is calculated based on the noise level. (2) The consideration of the coordinates of the sound source and the receiver will introduce fluctuating direct sound energy, which is very likely to make the predicted vocal effort more discrete.

2.2 Measurement site

Two typical Chinese college canteens were selected as the measurement site. One was the student canteen in the west area of Beichen Campus of Hebei University of Technology (H), and the other was the third student canteen in Weijin Road Campus of Tianjin University (T). Their interiors and room information were shown in Figure 1 and Table 1. The canteen H was a typical flat space, and the canteen T had a small mezzanine. In terms of the number of seats, these two canteens are both between 500 and 1,000, which are of the general level of the canteens of Chinese colleges [1] and have better universal applicability. There was little sound absorption in both of the two canteens. The floor materials of both canteens were ceramic tiles, and the surface of the walls was painted and wallpaper. The dining table was made of plastic and stainless steel. The ceiling material of the canteen H was gypsum board. The canteen T had no suspended ceiling, and the roof material was steel. In winter, there would be a small cushion on the chair, which was the rare sound-absorbing material in these two canteens except for people.

a b

Figure 1 Photos of the two canteens(a: canteen H; b: canteen T)

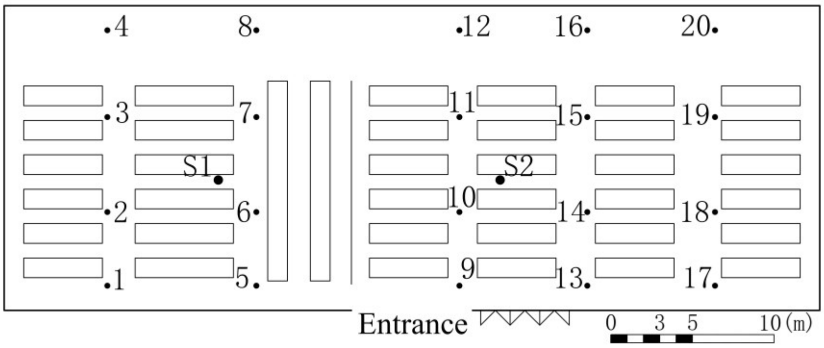
Table 1 Room information of two selected college canteens

	Length/Width /Height(m)	Volume (m^3)	Surface area (m^2)	Floor area (m^2)	No. of seats	Seating density ($1/m^2$)	RT (s)
H	49.8/18.6/3.4	3149.4	2317.7	926.3	680	0.73	1.08
T	47.5/23.4/8.9	8417.8	3184.5	1171.4	716	0.61	1.04

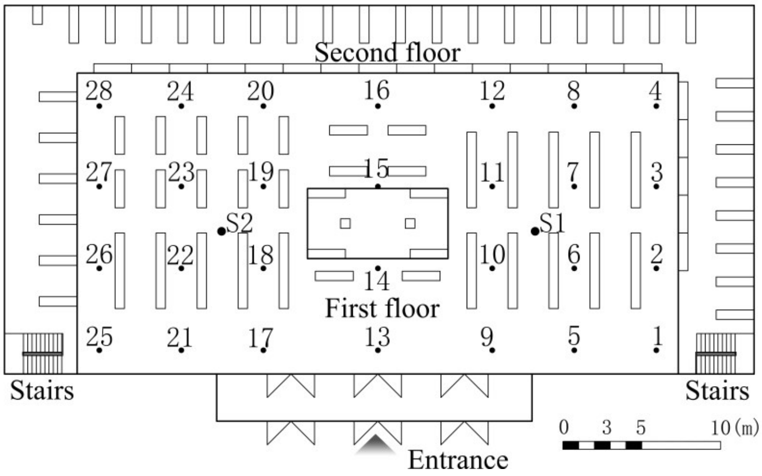
Reverberation time (RT) was measured by the impulse response method using Dirac v6.05 software based on ISO-3382 [34]. In order to avoid panic caused by the large impulse sound during the test, and the fire system's reaction caused by the gun smoke, the balloon burst was used as the sound source, which was shown as S1 and S2 in Figure 2. In canteen H, point 6 and 7 were the receivers of S1 and point 10 and 11 were the receivers of S2. In canteen T, point 10 and 11 were the receivers of S1 and point 18 and 19 were the receivers of S2. The reverberation time in this article was mainly used to predict the vocal power level from the measured noise sound pressure level, using the classical formula based on the assumption of the reverberant sound field. Considering that the noise level was A-weighted, it is reasonable to use the average value of reverberation time in several frequency bands. According to Barron's prediction model [35], 500 *Hz*, 1000 *Hz* and 2000 *Hz* were chosen as the averaged octave bands. The single value of reverberation time was obtained by averaging the values at the centre frequency of the three-octave bands (500 Hz, 1000 Hz, 2000 Hz). During the measurement of the reverberation time, there were few people in the two canteens. The results were shown in Table 1.

2.3 Noise level

The change of noise level with time was obtained by measurements at different receivers at 5-minutes intervals. The receivers were distributed evenly in the two canteens, as shown in Figure 2. The measurements lasted for three hours, from 11:00 to 14:00 in canteen H and from 10:45 to 13:45 in canteen T. Both measurements included a complete meal cycle (from a small number of occupants to the peak period, and then to a small number of occupants again). The hand-held sound level meters (AWA5688) were used, and the measurement period at each point was one minute.



a



b

Figure 2 Plans of the two canteens (a: canteen H; b: canteen T. S1 and S2 were the sound source in the measurements of reverberation time in two canteens, and the numbers were the receivers.)

2.4 The number of occupants in the canteens

In the two canteens, cameras were set at the entrance to record the number of occupants' entering and leaving during mealtime. And then, the change in the number of occupants with time was counted through the video records.

2.5 Speaking ratio

The status of diners was recorded by another camera. Diners facing the camera were numbered, and their speaking time and non-speaking time were counted using a small python program, as shown in Figure 3. When a person starts to eat, the program starts to run and automatically counts. When a diner starts talking, click a button, and when this diner stop talking, click another button. When the meal is over, the program automatically counts the mealtime and speaking ratio of the diners. Then, the speaking ratio was obtained by the statistics of everyone's speaking time and total mealtime. In canteen T, 229 diners were counted by video records.

Figure 3 Video screenshots of diners in the canteen T, in which the blue numbers represented the single diners, while the red ones were in a larger group.

2.6 Limitation

There are always some staff during the use of the canteen, and the noise of these people is regarded as background noise. This will result in higher background noise, which may introduce some uncertainties. In order to avoid causing panic among users who are always present in the college canteens, balloon burst, instead of gunshot, was used to obtain the impulse response, which is also a method allowed by ISO 3382.

Since this study was only carried out in two community college canteens, the stability of the data obtained needs more research to confirm.

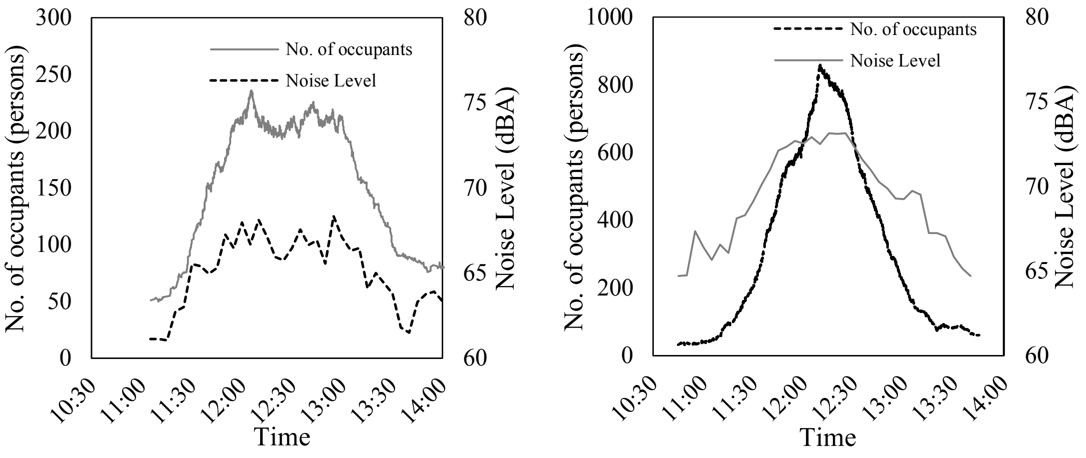
3 Results

3.1 Noise level and the number of occupants

The changes in the number of occupants and noise level during mealtime in the two college canteens were shown in Figures 4a and 4b (grey solid lines). The relationship between the number of occupants and time in the two canteens showed three stages: growth, plateau, and elimination. Since the two canteens were open almost all day, the canteens were always maintained with a small number of occupants in the non-dinner time. In canteen H, starting from 11 o'clock, the number of occupants had gradually increased from about 50 people to about 210 people at 12:00. Then, the number of occupants entered a peak period that lasted about an hour. From 13:00, the number of occupants began to decrease with a slope that was almost the same as increase, until 14:00 the number of occupants decreased to slightly more than the state at the beginning. The increase and decrease curves of the number of occupants had a good symmetry.

The relationship between the noise level and the time in the two canteens was also shown in

Figure 4. The noise-level curves were almost synchronised with the number of occupants in two canteens. However, near the peak of the number of occupants, the noise level appeared to reach a plateau, and the noise no longer increased with the increase in the number of occupants. There might be two reasons. The first was that after the noise reached a relatively high level, the increase in human voice gradually reached the physiological limit, and the proportional increase in the Lombard effect was suppressed, thereby reducing the increase in noise. The second reason might be that after the noise increased, the diner's willingness to speak was suppressed, which meant that the speaking ratio decreased.



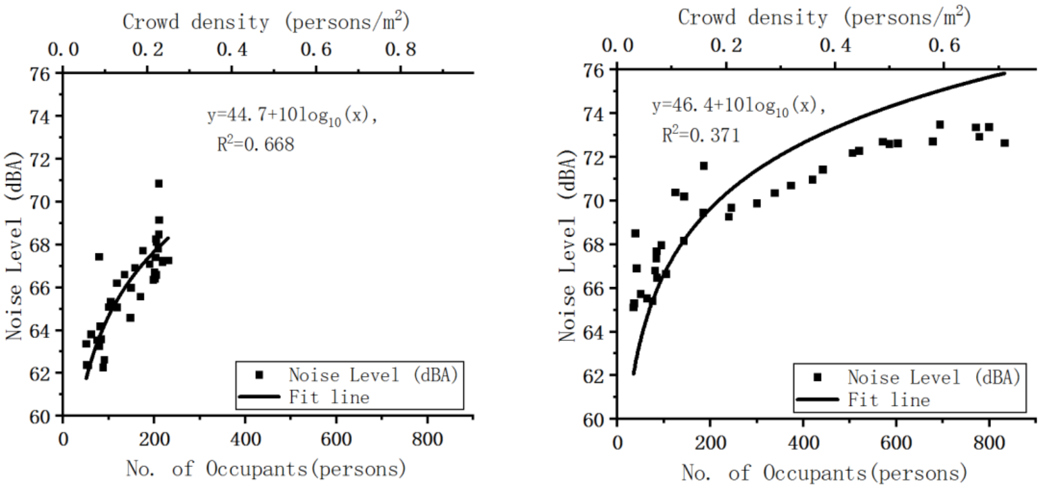
a b

Figure 4 The variation of the number of occupants and noise level during mealtime in the two canteens. (a: canteen H; b: canteen T. The measurement time is three hours, from 11:00 to 14:00 in canteen H and from 10:45 to 13:45 in canteen T. The solid grey line is the number of occupants, and the dashed black line is the noise level.)

The relationship between noise and the number of occupants in the two canteens was obtained, as shown in Figure 5. It showed that about fifty diners presented in the two canteens in the initial state so that the initial background noise was relatively high (62 ~ 65 dBA). The range of noise changing in both canteens is 8-10 dB, and the noise level had a strong correlation with the number of occupants. If the diner was regarded as an omni-source and each source produced the same contribution at the receiving point, then the relationship between the total sound pressure level at the receiving point and the number of diners can be simply described as.

$$SPL_{total\_omni} = SPL_{omni\_each} + 10 * \log_{10} N \qquad \text{Equation 10}$$

where, SPL_{total_omni} is the total sound pressure level at the receiving point; SPL_{omni_each} is the contribution of each source; N is the number of diners in the canteen. The fit line of $y=a+\log_{10}(x)$ was also shown as the solid line in Figure 5. It showed that the line fitted well with the scattering points in canteen H, with a coefficient of determination of 0.668. In canteen T, there was some deviation between the line and the scattering points, with a coefficient of determination of 0.371. So, the relationship between the noise level and the number of occupants was not simply the superposition of equal-intensity sound sources, and it depended on different canteens. Crowd density [5] was also an important factor affecting crowd noise, as shown in Figure 5. However, considering that the two halls were close in the plan area, the reason for the different noise level under the same number of people or density might be more dependent on some other differences between the two canteens. First, the maximum number of occupants in the canteen T was significantly higher than that in the canteen H, reaching about 800 occupants. Second, the change in the number of occupants in the canteen T did not have a flat peak period. It began to decline immediately after reaching the maximum number of occupants. The increase and decrease in the number of occupants also had a good symmetry, and the peak appears at about 12:10.



a b

Figure 5 The relationship between noise level and the number of occupants in the two canteens. (a: canteen H; b: canteen T. The solid line is the fit line of Equation 10, $y=a+\log_{10}(x)$.)

3.2 The speaking ratio

The relationship between speaking ratio and time was obtained by recording each diner during the mealtime, as shown in Figure 6. The results showed that the average speaking ratio varied from 0.06 to 0.26 during the meal. The average value of the speaking ratios was 0.12, which was far less than the practical value, 1/3.

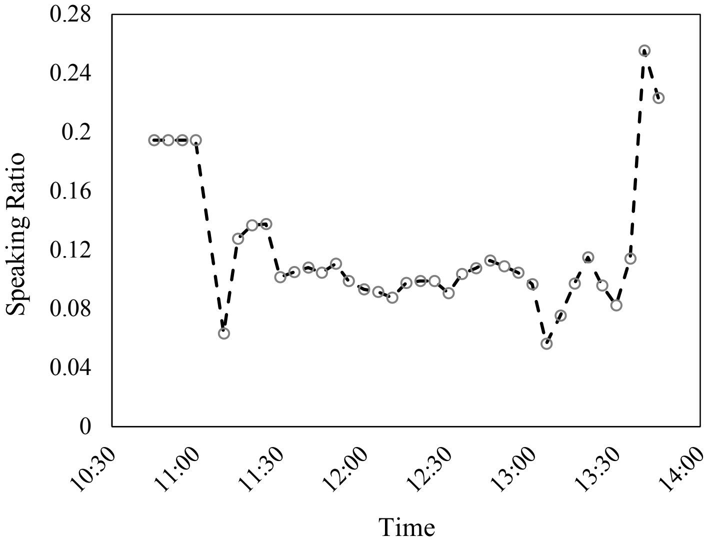


Figure 6 The relationship between speaking ratio and time in the Canteen T.

According to the one-to-one correspondence between the noise level and time, the relationship between the speaking ratio and the noise level was obtained, as shown in Figure 7. With the increase of noise level, the average value of the speaking ratio decreased significantly, from the initial about 0.2 to about 0.09, which indicated that when the environment became noisier, diners' desire to communicate would gradually weaken. Moreover, the variation of speaking ratio was high when the noise level was low. As the noise increased, it gradually converged, and the speaking ratio would not continue to decline and became stable at around 0.1 after the noise reached 70 dBA. The possible reason was that the initial increase in noise might suppress non-functional communication. After the noise level increased to a certain high value, the diners would only carry out the necessary functional communication, so the speaking ratio became stable and did not decrease. The discreteness of the speaking ratio data indicated that the speaking ratio has great individual differences.

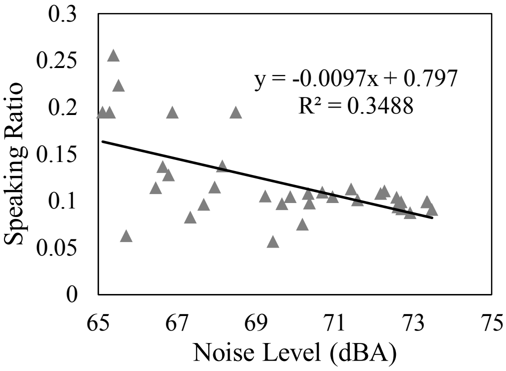


Figure 7 The variation of speaking ratio with the noise level. (the solid line is the linear regression.)

The relationship between the noise and speaking ratio of men and women were shown in Figure 8. The gender difference was obvious and varied with the noise level. The speaking ratio of men was higher than that of women when the noise was low. As the noise increased, the ratio of men's speaking began to decrease. When the noise reached 70 dB, the gender difference was no longer apparent, and the speaking ratio of men and women were both about 0.1.

In addition to the noise level and gender, group size was also an essential factor in the speaking ratio. It was generally believed that when there were a large number of groups, there was only one speaker at a time, and the other members were all listeners. According to this theory, as the group size increased, the speaking rate would show a downward trend. The relationship between the speaking ratio and the group size was shown in Figure 9. As the group size increased, the speaking ratio showed a downward trend. However, the speaking ratio in the three-person group was higher than that in the two-person group. The reason might be that the conversation was more heated after the group became larger, which caused the speaking ratio to rise. However, the speaking ratio in the seven-person group was also higher than that in the six-person group. Considering that the sample number of the seven-person group is only 3, this might also be caused by the fluctuation of the sample. If more samples were measured, this trend might disappear.

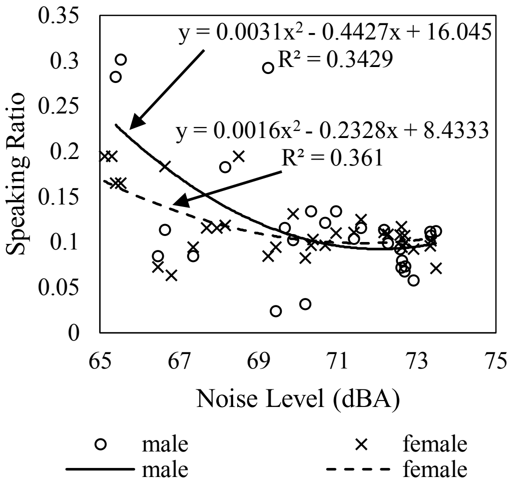


Figure 8 The difference in the speaking ratio in gender. (the male was represented by empty circles, and the female was represented by crosses. Solid lines represent polynomial regression lines for the male, and dashed lines represent polynomial regression lines for the female.)

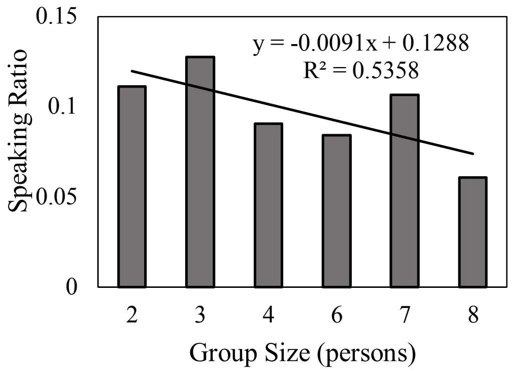


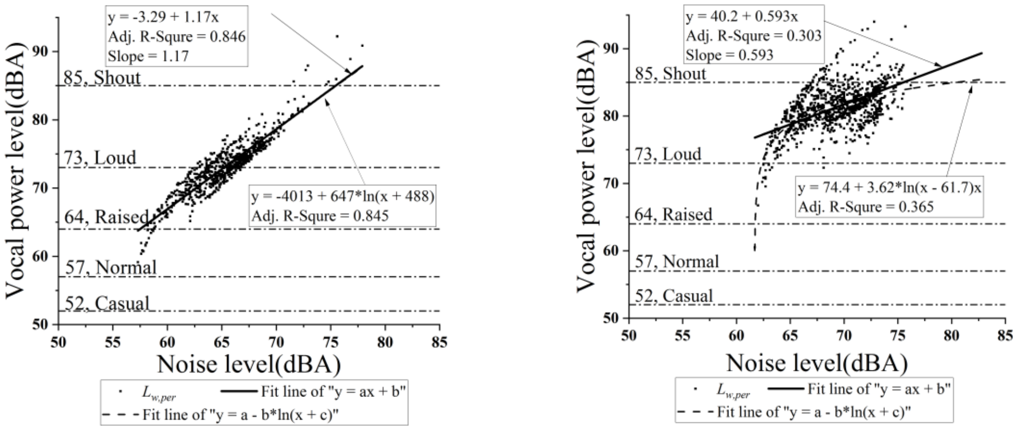
Figure 9 The relationship between speaking ratio and group size. (the solid line is the linear regression.)

From the above analysis, it indicated that the speaking ratio was not a fixed value. It also changed with noise and other factors, such as gender and group size. However, the deeper mechanism of this phenomenon was also not verified, and more research needs to be conducted to get a more in-depth explanation.

3.3 Lombard effect and evaluation of vocal effort

According to the one-to-one correspondence between the noise level and time, the relationship between the vocal power level $L_{W,speech}$ and the noise level in two canteens was also obtained respectively, as shown in Figure 10. Linear regression was used to find the Lombard slope, and two values of Lombard slope were obtained in the two canteens. The Lombard slope was 1.26 dB/dB in the canteen H and 0.80 dB/dB in the canteen T. The Lombard slopes in Chinese college canteen were higher than the results in similar non-laboratory tests as follows. In a series of non-laboratory conditions tests led by Pearson [20], the Lombard slope was 0.58 dB/dB in the home and 0.76 dB/dB in School. In the test in a 108 m^3 room conducted by Nijs [16], the maximum Lombard slope was 0.48 dB/dB. In a series of tests conducted by Hodgson in 13 restaurants, the overall Lombard slope was 0.69 dB/dB. A possible explanation for this was the cycling effect described in the Introduction. When a person entered the space, the increase in noise went through an iterative process, which made the increased noise level higher than that in the anechoic room. Another possible explanation for this was the cultural difference. Previous studies[5] had found that Chinese people showed higher sensitivity to different noises, which might appear as a higher Lombard slope.

It should be noted that the Lombard slope in this paper was obtained by linear regression, and its determination coefficient was not high, as shown in Figure 10. The relationship between noise and the number of people was very discrete in field tests[27, 36], with a coefficient of determination ranging from 0.05 to 0.85, which was also reflected in the subsequent relationship between noise and vocal effort. This looked more like a feature of crowd noise, the randomness of noise intensity and vocal time made it difficult to make predictions accurate to a certain value. The nonlinear relationship between voice and noise has also been introduced [27]. But due to its complexity, it was rarely used to find the Lombard slope, and simpler and easier-to-use algorithms were needed. Logarithmic regression was also performed in the results of the two halls, and a better coefficient of determination was obtained. In canteen H, the adjusted R-squared was almost the same, while in canteen T, the adjusted R-squared rised from 0.303 in linear regression to 0.365 in logarithmic regression, which indicated some non-linearity in the relationship between the noise level and vocal effort.



a b

Figure 10 The relationship between vocal power level and the noise level in two canteens. (a: canteen H; b: canteen T. The solid line is the linear regression. The dashed line is the logarithmic regression. The chain line is the Pearson's evaluation criterion (the former is the vocal power level, and the latter is the evaluation value.).)

The voice effort of people in the two canteens was at a high level. Most values of vocal power level in the canteen H were between 70 to 80 dBA, which was between Raised and Loud based on the Pearson's evaluation criterion, and most values of vocal power level in the canteen T varied from 75 to 85 dBA, which was between Loud and Shout. It could be seen that the noise in the Chinese college canteens was very high. In this environment, diners needed to increase the vocal power level to about 64 dBA (Raised) to make themselves be heard. In most cases, it needed to reach 73 dBA (Loud), and during peak dining hours, more than 85 dB (Shout) was required to maintain the communication.

The overall relationship between vocal power level $L_{W,lombard}$ and the noise level in two canteens was shown in Figure 11. The overall Lombard slope was 1.33 dB/dB, which was larger than the value in either canteen. 17.1% of the data fell between Raised and Loud, and 71.8% of the data fell between Loud and Shout. It indicated that the crowd noise in Chinese college canteens had a significant impact on communication and some acoustic treatment was necessary.

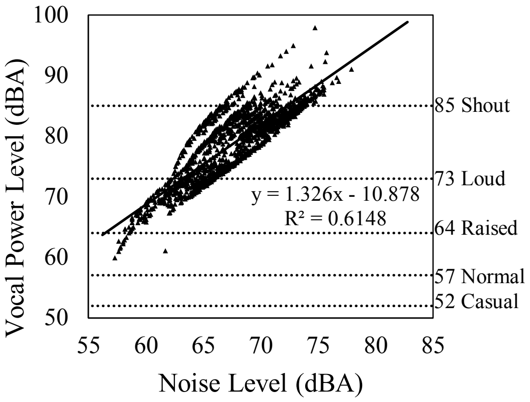


Figure 11 The variation of the vocal power level with the noise level in both canteens. (The solid line is the linear regression. The dashed line is the Pearson's evaluation criterion (the former is the vocal power level, and the latter is the evaluation value).)

4 Discussion

In the test of speaking ratio in this paper, results showed that the speaking ratio in Chinese college canteens during the peak dining period was about 0.1, which was far less than 1/3 of the experience value. 98.7% of values were less than 1/3. The reason why the results of the speaking ratio in this article were different from that in the previous studies might be the cultural difference. Chinese college canteen tended to focus more on dining functions, and the dining time was usually very short. There was not much time for communication during mealtime.

Secondly, it may be due to the dining environment. The two canteens accommodated a large number of occupants, which resulted in a smaller distance between diners and led to higher received noise and a lower speaking ratio. Moreover, the speaking ratio used in previous studies, 1/3, was a practical value rather than the result directly confirmed by testing. From this perspective, the actual value of the speaking ratio in the English environment also required more tests to verify.

5 Conclusion

In this paper, the relationship between the crowd noise and time was obtained in two typical Chinese college canteens. One was the western canteen of the Beichen campus of Hebei University of Technology, and the other was the third student canteen of the Tianjin University Weijin Road campus. During the meal, the noise varied from 61 dBA to 73 dBA, and the noise-time curve was inverted U-shaped. The maximum value of noise appeared at the peak of the meal from 12:10 to 12:30. The noise level had a strong correlation with the number of occupants, but the relationship between the two was not simply the superposition of equal-intensity sound sources.

As an important parameter of crowd noise, the speaking ratio was also studied in this paper by analysing the video recording. It varied from 0.06 to 0.26 during the meal. Noise seemed to be the most critical factor influencing the speaking ratio. The variation of speaking ratio was high when the background noise was low. As the noise increased, it gradually converged, while the speaking ratio did not continue to decline and became stable at around 0.1 after the noise reached 70 dBA. Gender and group size could also affect the speaking ratio. Based on the results in this paper, the average value of the speaking ratios was 0.12, which was far less than the practical value, 1/3.

A prediction model for crowd noise under crowd noise was introduced into Chinese college canteens, which considered direct and reverberant sound energy using the parameters of room information, the location of the diners, and the speaking ratio. The sound power level of diners was obtained by this model, and the Lombard slope was also obtained in the two canteens. The overall Lombard slope was 1.33 dB/dB. Based on Pearson's evaluation criterion, 17.1% of the data fell between Raised and Loud, and 71.8% of the data fell between Loud and Shout. It indicated that the crowd noise in Chinese college canteens had a significant impact on communication and some acoustic treatment was necessary.

Acknowledgements

This research was supported by a grant from the key projects of the scientific research program of colleges and universities in Hebei Province, China (No. ZD2017218).

Reference

[1] X. Zheng and S. Zhang, “Survey of acoustic environment of public canteens of universities,” *Diansheng Jishu(Audio Engineering)*, vol. 37, no. 2, pp. 11–13, 2013.

[2] W. To and A. Chung, “Restaurant noise: levels and temporal characteristics,” *Noise & Vibration Worldwide*, vol. 46, no. 8, pp. 11–17, 2015.

[3] G. Marsico, G. Brambilla, S. Curcuruto, M. Clapiz, R. Betti, and M. Riccardi, “Acoustic climate inside a canteen and mitigation solutions,” *Journal of the Acoustical Society of America*, vol. 123, no. 5, p. 3354, 2008.

[4] X. Chen and J. Kang, “Acoustic comfort in large dining spaces,” *Applied Acoustics*, vol. 115, pp. 166–172, 2017.

[5] Q. Meng, S. Zhang, and J. Kang, “Effects of typical dining styles on conversation behaviours and acoustic perception in restaurants in china,” *Building and Environment*, vol. 121, pp. 148–157, 2017.

[6] Z. Wang, “Analysis and prevention countermeasures of acoustic environment quality in college restaurant,” *Journal of Henan Mechanical and Electrical Engineering College*, vol. 20,

no. 4, pp. 42–44, 2012.

[7] X. He, "Evaluation method of the university canteen environmental comfort a case study of guangdong vocational college of environmental protection engineering," Master's thesis, South China University of Technology, 2013.

[8] A. White, "The effect of the building environment on occupants: the acoustics of dining spaces," *Robinson College*, 1999.

[9] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta Acustica united with Acustica*, vol. 86, no. 1, pp. 117–128, 2000.

[10] H. L. Pick Jr, G. M. Siegel, P. W. Fox, S. R. Garber, and J. K. Kearney, "Inhibiting the lombard effect," *The Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 894–900, 1989.

[11] M. Cooke, S. King, M. Garnier, and V. Aubanel, "The listening talker: A review of human and algorithmic context-induced modifications of speech," *Computer Speech & Language*, vol. 28, no. 2, pp. 543–571, 2014.

[12] S. A. Zollinger and H. Brumm, "The evolution of the lombard effect: 100 years of psychoacoustic research," *Behaviour*, vol. 148, no. 11-13, pp. 1173–1198, 2011.

[13] J.-C. Junqua, "The lombard reflex and its role on human listeners and automatic speech recognizers," *The Journal of the Acoustical Society of America*, vol. 93, no. 1, pp. 510–524, 1993.

[14] M. Garnier, N. Henrich, and D. Dubois, "Influence of sound immersion and communicative interaction on the lombard effect," *Journal of Speech, Language, and Hearing Research*, 2010.

[15] P. Bottalico, I. I. Passione, S. Graetzer, and E. J. Hunter, "Evaluation of the starting point of the lombard effect," *Acta Acustica United With Acustica*, vol. 103, no. 1, pp. 169–172, 2017.

[16] L. Nijs, K. Saher, and D. den Ouden, "Effect of room absorption on human vocal output in multitalker situations," *The Journal of the Acoustical Society of America*, vol. 123, no. 2, pp. 803–813, 2008.

[17] H. Lazarus, "Prediction of verbal communication is noise - a review: Part 1," *Applied Acoustics*, vol. 19, no. 6, pp. 439–464, 1986.

[18] M. D. Skowronski and J. G. Harris, "Applied principles of clear and lombard speech for automated intelligibility enhancement in noisy environments," *Speech Communication*, vol. 48, no. 5, pp. 549–558, 2006.

[19] J. Brunskog, A. C. Gade, G. P. Bellester, and L. R. Calbo, "Increase in voice level and speaker comfort in lecture rooms," *The Journal of the Acoustical Society of America*, vol. 125, no. 4, pp. 2072–2082, 2009.

[20] K. S. Pearsons, R. L. Bennett, and S. Fidell, "Speech levels in various noise environments," Tech. Rep., 1977.

[21] M. Hayne, J. Taylor, R. Rumble, and D. Mee, “Prediction of noise from small to medium sized crowds,” *Proceedings of Acoustics 2011*, 2011.

[22] H. Levitt and J. C. Webster, “Effects of noise and reverberation on speech,” *Handbook of acoustical measurements and noise control. McGraw-Hill, New York*, 1991.

[23] J. H. Rindel, “Verbal communication and noise in eating establishments,” *Applied Acoustics*, vol. 71, no. 12, pp. 1156–1161, 2010.

[24] M. Hayne, R. Rumble, and D. Mee, “Prediction of crowd noise,” in *Proceedings of the First Australasian Acoustical Societies Conference*, 2006.

[25] M. P. N. Navarro and R. L. Pimentel, “Speech interference in food courts of shopping centres,” *Applied acoustics*, vol. 68, no. 3, pp. 364–375, 2007.

[26] I. R. Cushing, F. F. Li, T. J. Cox, K. Worrall, and T. Jackson, “Vocal effort levels in anechoic conditions,” *Applied Acoustics*, vol. 72, no. 9, pp. 695 – 701, 2011.

[27] M. Hodgson, G. Steininger, and Z. Razavi, “Measurement and prediction of speech and noise levels and the lombard effect in eating establishments,” *The Journal of the Acoustical Society of America*, vol. 121, no. 4, pp. 2023–2033, 2007.

[28] J. H. Rindel, “Restaurant acoustics–the science behind verbal communication in eating establishments,” 2017.

[29] S. Tang, D. W. Chan, and K. Chan, “Prediction of sound-pressure level in an occupied enclosure,” *The Journal of the Acoustical Society of America*, vol. 101, no. 5, pp. 2990–2993, 1997.

[30] E. P. J. de Ruiter, “Lombard effect, speech communication and the design of large (public) spaces.” in *Forum Acusticum*, 2011.

[31] M. E. Beckman, “Prosodic structure and tempo in a sonority model of articulatory dynamics,” *Papers in laboratory phonology II: Segment, gesture, prosody*, 1992.

[32] J. Kang, “Comparison of speech intelligibility between english and chinese,” *The Journal of the Acoustical Society of America*, vol. 103, no. 2, pp. 1213–1216, 1998.

[33]H. Xie, H. Li, and J. Kang, “The characteristics and control strategies of aircraft noise in china,” *Applied Acoustics*, vol. 84, pp. 47–57, 2014.

[34] E. ISO, “3382-2, 2008, "acoustics - measurement of room acoustic parameters - part 2: Reverberation time in ordinary rooms",” *International Organization for Standardization, Brussels, Belgium*.

[35] M. Barron, “Theory and measurement of early, late and total sound levels in rooms,” *The Journal of the Acoustical Society of America*, vol. 137, no. 6, pp. 3087–3098, 2015.

[36] L. Zelem, V. Chmelk, D. Urbán, and M. Rychtáriková, “Analysis of the acoustic conditions in the student restaurant,” in *proc of Euronoise*, vol. 31, 2015.