Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 78 (2015) 3102 – 3107

Energy Procedia

6th International Building Physics Conference, IBPC 2015

Teachers and Teaching: Speech Production Accommodations Due to Changes in the Acoustic Environment

Eric J. Hunter[a]*, Pasquale Bottalico[a], Simone Graetzer[a], Timothy W. Leishman[b], Mark L. Berardi[b], Nathan G. Eyring[b], Zachary R. Jensen[b], Michael K. Rolins[b], Jennifer K. Whiting[b]

[a]*Communicative Sciences and Disorders, Michigan State University, 1026 Red Cedar Road, East Lansing, MI 48824*
[b]*Physics and Astronomy, Brigham Young University, N283 ESC, Provo, UT 84602*

Abstract

School teachers have an elevated risk of voice problems. The results of three studies investigating this elevated risk in terms of speaking environments teachers' encounter at work are presented. In the first study, 57 teachers were observed for 2 weeks (waking hours), comparing voice use in the school environment and in non-school environments. In a second study, 45 participants performed a short vocal task multiple times in two different rooms: a variable acoustic room and an anechoic chamber. Each time they entered the variable acoustics room, the reverberation time and/or the background noise condition had been modified. Besides being having their speech recorded, subjects responded to questions about their vocal comfort and their perception of changes in the acoustic environment. In a third study, 20 untrained vocalists performed a simple vocal task in the following conditions: with and without background babble and with and without transparent plexiglass shields to increase the first reflection. Relationships were examined between [1] the results for the room acoustic parameters; [2] the subjects' perception of the room; and [3] the recorded speech acoustic. From these three studies several gender differences were found; some of those differences held for each room condition (reverberation, noise level, and early reflection).

Peer-review under responsibility of the CENTRO CONGRESSI INTERNAZIONALE SRL

Keywords: aural feedback; speech production; Lombard Effect

* Corresponding author. Tel.: +1-517-884-67778; fax: +1-517-353-3176.
E-mail address: ejhunter@msu.edu

Peer-review under responsibility of the CENTRO CONGRESSI INTERNAZIONALE SRL
doi:10.1016/j.egypro.2015.11.764

1. Introduction

More than 18% of the primary and secondary school teachers in the US miss at least one day of work per year due to voice disorders (1). The individual and societal economic impact of voice problems experienced by teachers is significant. For example, teachers comprise only 4.2% of the U.S. workforce and 16% of the population of professional voice users. These voice problems affect not only the teachers, but several studies suggest that students with a teacher with a voice impairment are less likely to learn the information taught in class (2,3). Thus, previous studies have focused on understanding the causes of teachers' increased vocal risk. Frequently cited causes include the need to speak loudly over long periods of time at an elevated speaking fundamental frequency (4). Previous work has included full-day ambulatory monitoring of voice use (5, 6) and perceptual monitoring of fatigue (7).

Less frequently explored are potential environmental causes, including allergens or other airborne particulates, humidity level, and noise level (both student babble and environmental noise). Poor acoustics (e.g., high reverberation time, or extra low reverberation time) may also be a significant factor (8). While classrooms are often designed or retrofitted with vocal communication in mind (e.g., speech transmission index), it is possible that a classroom designed for intelligibility and speech transmission may be in fact detrimental to a teacher trying to teach. If true, this would add another factor to consider when designing classrooms so as to include not just speech conveyance but to consider the teacher (e.g. perception and speaking goals).

Studies suggest that variations in environment or communication context cause speakers to adjust their vocal behaviors (often sub-consciously). If these adjustments, often called speech accommodations or strategies, are unhealthy, increased vocal problems are likely. Within a classroom specifically, building design and changes in the acoustic environment, such as variance in noise and room reverberation, may trigger speech accommodations in teachers. For example, a frequently discussed accommodation is the Lombard effect, which describes a response to increased noise: females compensating with loudness, and males with fundamental frequency (9).

While many studies have considered classroom acoustics, few have looked at a classroom from the talker's perspective (10,11). Thus, the current studies examine the following research question: To what degree does acoustic environment and communication goal change how a talker produces speech in the environment? Towards answering this question, this manuscript summarizes three ongoing studies which contribute to different aspects of the problem. Each study addresses one aspect of the larger problem.

2. Study 1: Multiday observations

The purpose of the current study is to monitor a talker's voice use outside a laboratory and understand how they may change their speaking because of their environment.

2.1. Methods

In the first study, 57 teachers were observed for 2 weeks (waking hours) to compare how their voices changed in a school environment (mostly classroom) versus non-school environments (everything else). The primary resource for this study was the National Center for Voice and Speech (NCVS) teacher voice dosimetry databank, a databank containing two-week data blocks of speaking metric that have been captured as described previously (12,13). A short summary of the acquisition methods will be presented here for completeness.

The teachers, all K-12 teachers from more than a dozen schools in the Denver metropolitan area, Colorado, U.S.A., consisted of 45 females and 12 males with an average age of 44 (median, 55; s.d., 10). The subject breakdown by teaching grade was: K-4th grade, 59%; 5-8th grade, 16%; and 9-12th grade, 25%. Teacher breakdown by topic was: general classroom instruction, 71%; music/theater instruction, 16%; physical education instruction, 9%; and other (e.g., library instruction, special education), 4%. Voice dosimeter recordings were calibrated for each teacher's radiated speech level (14,15). Each teacher was taught how to attach and use the dosimeter, although a laboratory technician was on call at all hours to provide technical support. The device recorded data (speaking level in dB and speaking fundamental frequency in Hz) every 30 ms, with each data record time stamped so that the record could be searched by date and time for analysis. For each teacher who completed the two-week dosimetry study, a complete data set contained approximately 108,000 data records per hour, nearly 2 million records per day (18 hours), or 27 million records per 14-day period.

MATLAB scripts were written that could search all of the teacher dosimeter data by date and time, extracting fundamental frequency (F0) and speech level (dB). Using these scripts, average voicing measures were calculated in 15-minute increments throughout all the days. If at least 30 seconds of voicing were available within a 15-minute increment, the data were utilized for further statistical analysis. Data were compiled first into school and non-school environments based on the time of day: school environment (weekdays, 9am-2:30pm) and non-school environments (weekdays, 4:30pm to 10:00pm, and weekends). Treating each 15-minute increment average as one of many voice samples from a subject, linear mixed-effects models (fit by Maximum Likelihood) were implemented using R (www.r-project.org, ver 3.1.2, lme4), which were then used to compare the school and non-school values for both weekdays and weekends. This was done on F0 (in semitones), log(F0 in Hz) and dB; semitones and log(F0); semitones and log(F0) were used so that F0 values in the long recording had a more normal distribution. For the 57 teachers, data were collected from 769 days of 798 possible days, and usable voice data consisted of 8451 hours (weekdays: 6106 hrs; weekends: 2345 hrs).

2.2. *Results*

Figure 1 shows the averages of dB and F0 (in semitones), divided into school and non-school environments, for all subjects. Female and male teachers appeared to treat the two environments differently. Females' F0 was significantly lower for non-school environments (log(F0) was $p<0.05$, F0 in semitones was $p<0.005$). Not surprisingly, the difference between voicing produced in weekday and weekend non-school environments (during the same hours of the day) was not significant. Further, larger effects were shown in terms of the dB values, with the non-school weekday and weekend values significantly lower than the school values ($p<0.0001$). While the males teachers also showed a significant difference for school values compared to non-school values for dB changes ($p<0.0001$), the males raised their vocal pitch (log(F0) was $p<0.0001$, F0 in semitones was $p<0.01$). The statistical models also reveals another gender difference. For the males, there was no interaction of the time of day or weekday and weekend, yet for females there was significant interaction ($p<0.05$, $p<0.0005$, and $p<0.0001$ for log(F0), F0 in semitones, and dB respectively). This implies that in non-school environments, even after a day of work, the males behaved much the same in both weekdays and weekends, seemingly contradicting previous laboratory studies (9). Nevertheless, this gender difference may have implications on gender vocal health differences (16).

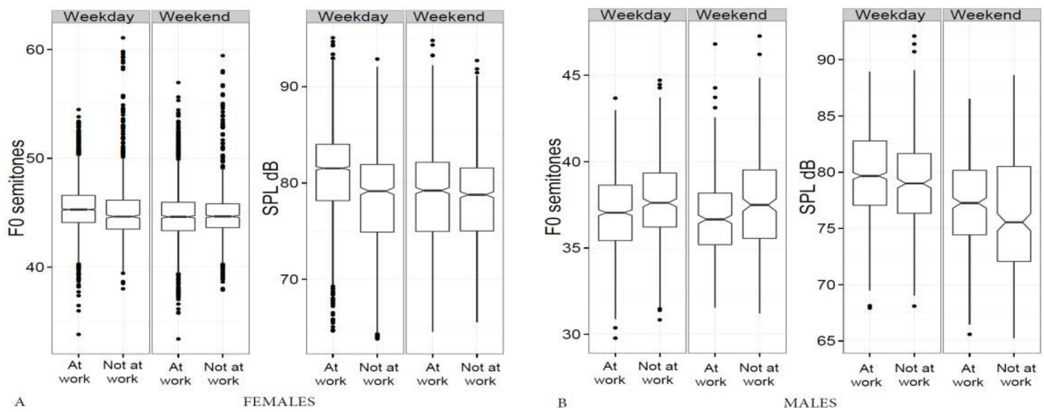

Fig. 1. Notched box plots for the male teachers (a) and female teachers (b) showing compiled times between 9am-2:30pm (school) and 4:30pm to 10:00pm (non-school).

3. Study 2: Variable Acoustics: Blinded

The purpose of this study was to test a subject's acuity to undisclosed changes in a variable acoustic chamber with the goal of understanding how differences in a room's acoustical characteristics may trigger different speech accommodation strategies.

3.1. Methods

Forty-five participants performed a short vocal task in two different rooms: a variable acoustic room and an anechoic chamber (Brigham Young University-Provo). Participants were equipped with a head-worn microphone and a neck-worn accelerometer/microphone. Subjects were instructed on proper speech elicitation protocol (e.g., using conversational voice, repeating mistaken speech tasks). Subjects were blinded to the study purpose; when each subject was first taken to the anechoic chamber, she/he was told that the study was intended to investigate how speech may change as a result of the unusual acoustical environment in the anechoic chamber. (This deception protocol was conducted after approval from the local Institutional Review Board.)

Undisclosed to the participants, three acoustical conditions were employed in the variable acoustics chamber: [1] low-level brown noise (sounding much like an HVAC system found in classrooms) with many sound-absorbing panels, [2] higher-level brown noise with many sound-absorbing panels, and [3] low-level brown noise with few sound-absorbing panels. The low-level noise at the talker position was approximately 35 dBA, whereas the higher-level noise was about 50 dBA. With many absorption panels in place, the room reverberation time was about 0.2 seconds. With few panels in place, room reverberation time increased to approximately 0.5 seconds. To ensure that the presence or absence of panels was not visually apparent to the participants, the potential panel positions were masked by visually opaque grill cloth. The verification of the room was conducted according to ISO 3382-2 for survey accuracy. The impulse response was measured using two 1/2” microphones and a dodecahedron loudspeaker.

The subject was first recorded in the anechoic chamber, then in the variable acoustics chamber with a random selection of previously mentioned room conditions. Each was recorded again in the anechoic chamber, followed by another recording in the variable acoustics chamber under a different room condition. This process was repeated for a total of six recordings (three in the anechoic chamber and three in the variable acoustics chamber) so that each variable acoustics chamber permutation was included in the recording sequence. While several speech tasks were recorded, only the analysis of the second and third sentences of the rainbow passage are presented here.

3.2. Results

For the low noise condition, participants on average had a higher fundamental frequency (Δf=1.87 Hz) in the reverberant case (0.5 sec) compared to the lower reverberant case (0.2 sec). Most subjects did not significantly change their sound level between the two cases. However, subjects with questionnaire reports of dehydration did have a significant change in dB (0.61 dB, $p<0.026$) with changes in reverberation.

To help validate the analysis, the variation of the speech parameters between the three different recordings in the anechoic chamber were observed. Across these three recordings, there were no significant changes, suggesting that fatigue or familiarity with the speech tasks were not factors. Further, three tests were conducted to determine the sensitivity of trimming on the analysis: [1] the recordings of the entire rainbow passage; [2] the second and third sentences of the rainbow passage; and [3] small 5-second segments of the rainbow passage (achieved by extracting the section of the recording between the 3rd and 8th seconds. Since not every participant spoke at the same rate these segments were not identical). Upon comparison, similar trends were observed for the all trimmings of the passage.

4. Study 3: Effects of short-term reflections

The purpose of this study was to test the effects of early reflection on speech production. Specifically, by giving a talker an early reflection without changing reverberation time, does speech production change due to early reflections?

4.1. Methods

Ten male and ten female non-smoking participants (ages of 18-29) were recruited with self-reported normal speech and hearing. The following conditions were imposed on the classroom: [1] with and without two reflective panels placed at 45^0 from the axis (to the front, left and right), at a distance of 0.1m; and [2] with and without a children’s babble presented at 60dB (60 +/- 1.5dB) at the talker’s position. The subjects were asked to produce speech at three volumes (loud, normal, soft but not a whisper). The instructions for speaking and reading were:

- Loud: *Imagine you are in a classroom and you want to be heard by all of the children.*
- Normal: *Speak normally.*
- Soft/confidential: *Imagine you are saying something to a friend who is next to you. You want her to hear you but no one else. But don't whisper.*

Talkers were recorded reading three standard passages with a head-mounted microphone, an accelerometer on the neck, and a boundary reference microphone at two meters. The three passages read were the Marvin Williams passage, the first paragraph of the Rainbow passage, and the Stella passage. After each condition (e.g., reading the text with a given style and in a given acoustic condition), subjects answered three questions on a visual analog (10cm) scale from *not at all* to *extremely*:

- Fatigue: *How fatigued would your voice be if you were to speak continuously in this condition for 20 minutes?*
- Comfort: *How comfortable was it to speak in this condition?*
- Control: *How well were you able to control your voice in this condition?*

Experiments were conducted in a classroom at Michigan State University. For the classroom, room acoustics parameters were measured in an unoccupied state without furniture from the impulse responses generated by a balloon pop (following the ISO 3382-2:2008). Four source positions and three microphone positions were chosen in the room, where 12 impulse responses were recorded. Mean T20 values in the octave band ranging from 125 Hz to 8000 Hz, along with the just noticeable difference (JND), were calculated. The mean T20 for combined 500 Hz and 1 kHz octave bands is 0.53 s (s.d. 0.04), while the C50 is 5.7 dB (s.d. 1.2).

As far as the reverberation time (T20) is regarding, the standard deviation of the mean spatial values (s.d. 0.01) is lower than the JND (0.03 s) and therefore demonstrate rather uniform spatial behaviour. As far as the clarity (C50) is regarding, the values range between 3.52 dB to 7.47 dB and higher values were found in the positions closer to the window.

The reverberation time in the room is consistent with ANSI Standard S12.60:2010. In order to modify the clarity perceived by subjects, two reflective panels were placed at 45°, 1m from subject. The objective difference in early reflections (C50) was measured with KEMAR 45BB-1 (HATS) with Genelec 8010 loudspeaker as a mouth.

Analysis included participants responses to the questions and the within-subject centered SPL values ("SPL delta"), calculated from the recording by subtracting the mean of SPL values across conditions per subject from the SPL value for a given condition. A linear mixed-effects model fit by Maximum Likelihood was run with the response variable, SPL delta, the fixed factors, style, room and panel, and the random factors, subject and gender.

4.2. Results

As would be expected, there was a difference in the speaking style at about 5 dB from Confidential to Normal to Loud, similar to the ISO-9921 standard. The within-subject centered SPL values ("SPL delta") were derived from the head-mounted microphone signal by subtracting the mean of SPL values across conditions per subject from the SPL value for a given condition (Fig. 2). Not surprisingly, there were elevated speech levels with noise yet there was a tendency for subjects to decrease their speech level when panels are present. From the questionnaires, some evidence of the perception of the effects of noise, style and panel was quantified. Responses to the comfort and control question were positively correlated ($r = 0.66$), while fatigue, as expected, was negatively correlated with both ($r = -0.5$; $r = -0.34$, respectively). In the loud condition (as compared to the normal condition), subjects reported that comfort and control decreased, while fatigue increased. Relatedly, comfort and control were greatest when noise was absent, especially in the normal style. These inverse correlations between vocal comfort and fatigue, and between comfort and noise are consistent with previous studies (11). The fatigue question garnered the greatest response in the loud style, especially when noise was present. Confidential style was similar to normal style with regard to fatigue, but was associated with lower comfort and control responses than normal style. Responses for confidential style were generally more variable. When panels were present and for normal and loud speaking style (no noise condition only), fatigue tended to decrease. Further, in the loud style, comfort tended to increase when panels were present. In the loud style and during noise, control tended to increase for the panel condition. These results imply that even while keeping reverberation time the same, reflective surfaces could be optimized by teachers in classrooms to increase perceived voice comfort.

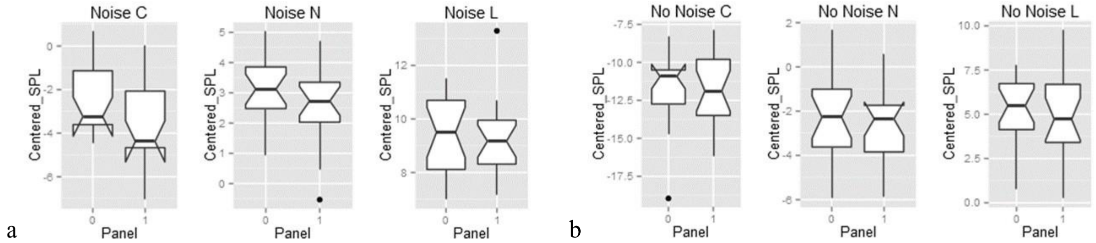

Fig. 2. Notched box plots for the noise (a) / no noise (b) condition. Confidential, Normal and Loud speaking styles are shown left /right.

5. Conclusions

With the elevated voice risk of occupations like school teachers, understanding the occupational environment and the vocal health risk factors has been a growing health interest. Fundamental to that interest is understanding the effect of a room on a talker. The studies presented here show the acuity talkers have to changes of acoustic environment which can be noted in a talker's speech, whether conscious of it or not. The results are not only relevant to room design but also to understanding talkers' perception of and speech adjustments to room acoustics.

Acknowledgements

Research reported in this publication in part supported by the National Institute on Deafness and Other Communication Disorders, Award Number R01DC012315. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

[1] Roy N, Merrill RM, Thibeault S, Gray SD, Smith EM. (2004). Voice disorders in teachers and the general population: effects on work performance, attendance, and future career choices. J Speech Lang Hear Res. 47(3):542-51.
[2] Morton V, Watson DR. (2001). The impact of impaired vocal quality on children's ability to process spoken language. Logoped Phoniatr Vocol. 26(1):17–25.
[3] Rogerson J, Dodd B. (2005). Is there an effect of dysphonic teacher's voices on children's processing of spoken language? J Voice. 19(1):47–60.
[4] Hunter EJ, Titze IR. (2009). Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise. Ann Otol Rhinol Laryngol. 118(6):449–60.
[5] Hunter EJ. (2012). Teacher Response to Ambulatory Monitoring of Voice. Logoped Phoniatr Vocol. 37(3):133-135.
[6] Gaskill CS, O'Brien SG, Tinter SR. (2012). The effect of voice amplification on occupational vocal dose in elementary school teachers. J Voice. 26(5):667.e19-27.
[7] Halpern AE, Spielman J, Hunter EJ, Titze IR. (2009). The inability to produce soft phonation (IPSV): A tool to detect vocal change in school teachers. Logoped Phoniatr Vocol. 28:1-11.
[8] Astolfi A, Pellerey F. (2008). Subjective and objective assessment of acoustical and overall environmental quality in secondary school classrooms. J Acoust Soc Am. 123(1):163.
[9] Letowski T, Frank T, Caravella J. (1993). Acoustical properties of speech produced in noise presented through supra-aural earphones. Ear Hear. 14(5):332–8.
[10] Pelegrín-García D, Fuentes-Mendizábal O, Brunskog J, Jeong C-H. (2011). Equal autophonic level curves under different room acoustics conditions. J Acoust Soc Am. 130(1):228–38.
[11] Pelegrín-García D, Brunskog J. (2012). Speakers' comfort and voice level variation in classrooms: laboratory research. J Acoust Soc Am. 132(1):249–60.
[12] Hunter EJ, Titze IR. (2010). Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. J Speech Lang Hear Res. 53(4):862–75.
[13] Titze IR, Hunter EJ, Svec JG. (2007). Voicing and silence periods in daily and weekly vocalizations of teachers. J Acoust Soc Am. 121(1): 469–78.
[14] Popolo PS, Svec JG, Titze IR. (2005). Adaptation of a Pocket PC for use as a wearable voice dosimeter. J Speech Lang Hear Res JSLHR. 48(4):780–91.
[15] Svec JG, Titze IR, Popolo PS. (2005). Estimation of sound pressure levels of voiced speech from skin vibration of the neck. J Acoust Soc Am. 117(3 Pt 1):1386–94.
[16] Hunter EJ, Tanner K, Smith ME. (2011). Gender differences affecting vocal health of women in vocally demanding careers. Logoped Phoniatr Vocol. 36(3-4):128–36.