Speech Prosody 2016
31 May - 3 June 2016, Boston, USA

Effect of Reverberation Time on Vocal Fatigue

Pasquale Bottalico, Simone Graetzer, Eric J. Hunter

Voice Biomechanics and Acoustics Laboratory
Department of Communicative Sciences and Disorders
Michigan State University
`pb@msu.edu, sgraetz@msu.edu, ejhunter@msu.edu`

Abstract

Vocal effort is a physiological magnitude that accounts for the changes in voice production that occur as vocal loading, which is the stress inflicted on the vocal folds when speaking for long periods, increases. It has been quantified in terms of Sound Pressure Level (SPL). In previous research, it has been shown that prolonged vocal effort can lead to vocal fatigue. An experiment was conducted to measure the effect of reverberation time on vocal fatigue, by means of an evaluation of variation in vocal effort over time. Twenty subjects were recorded while reading a text in anechoic, semi-reverberant and reverberant rooms in the presence of babble noise. Within-subject variation in SPL was measured per task. It was found that SPL tended to increase less over time in the semi-reverberant environment than in the anechoic or reverberant environments. This finding suggests that subjects experienced less short-term vocal fatigue in the semi-reverberant environment.

Index Terms: vocal effort, vocal fatigue, reverberation time, room acoustics

1. Introduction

In this study, the effects of room acoustics, voice style (corresponding to normal and raised levels) and vocal fatigue on vocal effort (quantified as Sound Pressure Level or SPL) and self-reported vocal effort, control, comfort and clarity, are considered.

The interaction between a person, a room and an activity leads to different sensations relating to voice production. This interaction determines the acoustic comfort, which contributes to human well-being. It also determines vocal comfort, which is related to all aspects that reduce vocal effort [1]. Vocal comfort appears to decrease with the speaker's perceived fatigue and the sensation of needing to increase the voice level [2]. The maximization of intelligibility, clarity, vocal comfort and control, and the minimization of vocal fatigue and effort, should be the priority of any professional talker. This is particularly important when the person is at elevated risk of vocal injury, such as in the case of teachers [3]. In classrooms, noise levels are typically high and acoustical conditions are not optimized for the talker but for the listener.

Vocal control can be defined as the capacity to self-regulate vocal behavior, e.g., Sound Pressure Level (SPL), and resonance. The sensation of control relates to the ability to adjust the voice consciously. In a communication environment, in general, speakers try to control their voice production in order to increase speech intelligibility. For example, while considering a communication partner with hearing limitations, a talker (deliberately or inadvertently) uses "clear speech", which has been characterized by a slower speech rate, a wider range of fundamental frequency (fo), and a higher temporal modulation index than conversational speech. Likewise, when talking in a noisy environment, people tend to raise the level of their voice (vocal effort) in order to maintain understandable communication [4].

Vocal effort is the exertion of the speaker as quantified by the A-weighted SPL (dB) at 1m distance from the mouth [5]. It is a physiological entity that accounts for changes in voice production when loading increases [6]. It relates to various factors such as the type of interlocutor, the speaker-listener distance, the background noise level, and other acoustic characteristics of the room [7, 8], linguistic factors such as vowel quality [8], and the speaker's level of fatigue 10, 11].

The effect of reverberation time and speaker-listener distance on voice power level was investigated by Pelegrín-García *et al.* [7]. Thirteen male talkers were recorded in four different environments: an anechoic chamber, a lecture hall, a corridor, and a reverberant room with average reverberation time (T30, 0.5–1 kHz) of 0.04 s, 1.88 s, 2.34 s and 5.38 s, respectively. The voice power level was found to depend almost linearly on the logarithm of the distance (with slopes between 1.3 and 2.2 dB per doubling distance) and changed significantly among rooms (intercepts between 54.8 and 56.8 dB). With the exception of the reverberant room, voice power level decreased as reverberation time increased.

Reverberation time was found to influence vocal intensity in continuous speech by Black [8]. He reported an analysis of SPL measured in the context of read speech produced by 23 males in 8 rooms differing in shape (rectangular and drum), for size (4.2 m^3 and 45.3 m^3), and reverberation time (0.2-0.3 s and 0.8-1.0 s). Greater vocal intensity was found in the dead (less reverberant) room than in the live rooms. Moreover, the speakers' intensity was found to be lower during the reading of successive phrases (comparing the first 3 and last 3 phrases of a total of 12 phrases) in live rooms than in dead rooms.

Voice power has been found to be related to room size and the magnitude of amplification by the room of the talker's voice at his/her ears, compared to anechoic conditions (termed "room gain" and "voice support"). Brunskog *et al.* [12] investigated objectively measurable parameters of the rooms related to any increase of the voice sound power produced by speakers and to the speakers' subjective judgments about six different rooms with different sizes, reverberation times and other physical attributes. They found that the speaker's voice power level when teaching is positively correlated with room size and inversely correlated with room gain.

10.21437/SpeechProsody.2016-101

Vocal fatigue is often experienced by speakers who use their voice for long periods and sometimes with increased vocal effort. Titze [13] identified two physiological aspects of such fatigue: laryngeal muscle fatigue and laryngeal tissue fatigue. Laryngeal muscle fatigue, which can involve tension in the vocal folds, appears to be caused by a depletion or accumulation of biochemical substances in the muscle fibers. Laryngeal tissue fatigue take place in non-muscular tissue layers (epithelium, superficial, and intermediate layers of the lamina propria) and appears to be caused by temporary changes in molecular structure that result from mechanical loading and unloading, *i.e.*, phonation [1]. The minimization of vocal fatigue is particularly important when (1) the speaker is at high risk of vocal injury, such as in teaching environments [14], when classroom acoustics are poor [15]; and (2) when vocal function is impaired by loading and/or incomplete muscle recovery [16].

Several studies have attempted to measure auditory perceptual and acoustic changes in the speech signal that are caused by prolonged and excessive vocalization. SPL has been found to be affected by vocal loading. Rantala *et al.* [10] analyzed recordings of 33 female teachers during the first and the last lessons (35-45min.) of a normal work day (5h). They divided the teachers into two categories: subjects who reported frequent symptoms of vocal fatigue, and subjects with few vocal complaints. SPL was found to increase by 0.5 dB between the first and last lesson, but this finding did not reach significance. Laukkanen *et al.* [11], who examined male teachers' voices before and after a working day with the same division of subjects into groups, found that SPL increased after the working day in both groups. Measurements of minimum attainable and comfortable vocal intensity were typically elevated following prolonged reading, as indicative of a rise in phonatory threshold. [17] Acoustic measures, such as jitter, shimmer, signal-to-noise ratio, showed a slight deterioration for untrained participants and remained relatively constant for participants with voice training. [18]

In summary, previous research suggests that the speech level can decrease under more reverberant conditions [7, 8], and increase as vocal fatigue increases [10, 11].

2. Experimental method

2.1. Subjects and conditions

The speech of 20 talkers (10 males and 10 females, who were non-smokers, with a mean age of 20.8y, with self-reported normal speech and hearing) were recorded in 3 different rooms in the presence of artificial babble noise. The noise was used in order to simulate noise present in a typical primary school classroom. The text was a 6 sentence excerpt from the Rainbow passage. Subjects were recorded while reading a text twice in two different speech styles (normal and loud) in three different acoustic environments (a total of 12 tasks per subject). The first was an anechoic room with dimensions 3.4×4.6×2.4m^3 (IAC 107840). The second was a semi-reverberant room, 8.5×7.3×4.6m^3. The walls were concrete block, the ceiling was concrete, and the floor was covered with vinyl tile. The third room was a reverberant room with dimensions 7.7×6.4×3.6m^3 (IAC 107840). The order of the 12 tasks was randomized in order to distribute possible vocal fatigue effects across all the tasks. Between tasks, the subjects took short breaks.

2.2. Equipment

Speech was recorded by a head-mounted omnidirectional microphone placed 5-7 cm from the mouth (Glottal Enterprises M80, Glottal Enterprises, U.S.A.). As this distance is lower than the critical distances in the three rooms, the SPL recorded by the microphone was associated only with the direct sound of the speakers. Speech was also recorded by an omnidirectional ultra-linear condenser microphone at 1m from the subject (ECM8000, Behringer, Germany). Microphones were connected to a PC via an external sound board (Scarlett 2i4, Focurite, U.S.A.). The signals were recorded with Audacity 2.0.6 with a sampling rate of 44100 Hz and processed to calculate measures of SPL.

2.3. Room acoustic parameters

Room acoustic parameters were obtained from the impulse response measurements in the non-occupied condition for the three rooms [18]. Balloon pops were used as impulses. The average reverberation time, T30, for combined 500 Hz and 1 kHz octave bands, were determined for each room in four different positions. The T30 for combined 500 Hz and 1 kHz octave bands was 0.04s (s.d. 0.005) in the anechoic room, 0.78s (s.d. 0.012) in the semi-reverberant room and 2.37s (s.d. 0.167) in the reverberant room. The measured values of the reverberation time for the three rooms between 125 and 8kHz are given in Table I.

During all the tasks performed by each subject, multi-talker children's babble was emitted by a directional loudspeaker placed at 2m in front of the subject. The power level of the loudspeaker was set in order to obtain an A-weighed equivalent level averaging both ears of 62dB in the talker position (measured with the Head and Torso Simulator).

Table 1. *Reverberation time and standard deviation per octave band in the three room.*

T30 (s.d.)	*Anechoic Room*	*Semi-reverberant Room*	*Reverberant Room*
125 Hz	0.05 (0.005)	1.01 (0.02)	1.26 (0.06)
250 Hz	0.04 (0.007)	0.92 (0.06)	1.82 (0.03)
500 Hz	0.04 (0.006)	0.78 (0.02)	2.23 (0.06)
1000 Hz	0.04 (0.005)	0.79 (0.01)	2.52 (0.04)
2000 Hz	0.04 (0.005)	0.79 (0.02)	2.33 (0.11)
4000 Hz	0.04 (0.005)	0.73 (0.02)	1.66 (0.05)
8000 Hz	0.04 (0.007)	0.55 (0.02)	1.01 (0.02)

2.4. Processing of the voice recording

Analysis of the SPL was performed with Matlab version 2014b. For each task, a time history with SPL evaluated at 0.125s intervals was obtained. Hence, 12 time histories were obtained per subject. The average over all of the SPL values was computed per subject, and this mean was subtracted from each value of the 12 time histories performed by that subject. This within-subject centering was performed in order to evaluate the variation in the subject's vocal behavior in the different conditions from the "mean" vocal behavior. After transformation, the parameter was termed ΔSPL.

2.5. Statistical method

Statistical analysis was conducted using R version 3.1.2. Linear mixed models (LMEs) were fit by REstricted

Maximum Likelihood (REML). Models were built using lme4, lmerTest and multcomp packages. Tukey's post-hoc pair-wise comparisons were performed to examine the differences between all levels of the fixed factors of interest. The p values for these multiple comparisons were adjusted using the single-step method. The LME output includes the estimates of the fixed effects coefficients, the standard error associated with the estimate, the degrees of freedom, df, the test statistic, t, and the p value. The Satterthwaite method was used to approximate degrees of freedom and calculate p values.

3. Results

A model was fit to the response variable ΔSPL (dB), with the fixed factors (1) room, (2) gender, and (3) chronological task order, with interactions of (4) room and order and two simple random effects terms (intercepts) for subject and time (measured in ms per task).

The estimates of the standard deviations of the random effects were 0.56dB for subject and 1.49dB for time. The residual standard deviation was 10.1dB. The estimate for the intercept was -0.002dB. The reference levels were semi-reverberant room and female gender. The estimate for the anechoic room was 1.37dB lower than that for the semi-reverberant room ($p<0.0001$), while the estimate for the reverberant room was 1.49dB lower than that in the semi-reverberant room ($p<0.0001$). Post-hoc comparisons confirmed the difference between the rooms, as shown in Figure 1.

The slope ΔSPL – chronological task order was observed to depend on room: 0.09dB/task in the semi-reverberant room ($p<0.01$), 0.28dB/task in the anechoic room ($p<0.0001$) and 0.21dB/task in the reverberant room ($p<0.01$). This effect is shown in Figures 2-4. Three simplified linear regression models were fit for the relationship between room and order, one per room, with the same random effects structure as the larger model. The best linear fits in anechoic, semi-reverberant and reverberant rooms are reported in equations (1), (2) and (3), respectively,

$$\Delta SPL_{anechoic} = -3.37 + 0.55 \cdot Order \tag{1}$$

$$\Delta SPL_{semi-reverberant} = 0.07 + 0.04 \cdot Order \tag{2}$$

$$\Delta SPL_{reverberant} = -0.88 + 0.08 \cdot Order \tag{3}$$

where *Order* represent the chronological order of task administration from 1 to 12. The effect of gender is not discussed here.

4. Discussion

An increase in SPL across the 12 tasks was observed, which may indicate that the subjects were experiencing vocal fatigue. This finding is consistent with the tendency for SPL to increase with vocal loading, as observed by, e.g., [10] and [11]. Overall, reverberation time and SPL were found to be inversely related such that as T30 increased from 0.8 to 2.37, there was a decrease in SPL of 0.73dB.

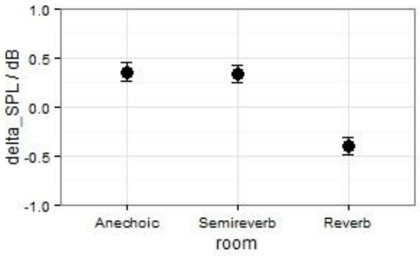
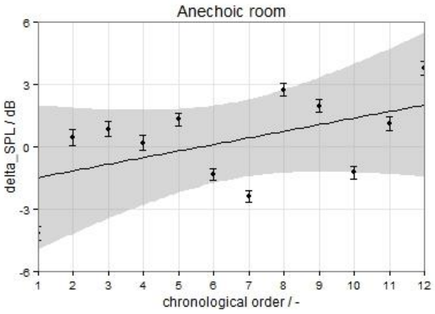


Figure 1 *Vocal effort (ΔSPL) in the three rooms.*



.

Figure 2 *ΔSPL by chronological task order with best fit linear regression lines in the anechoic room.*

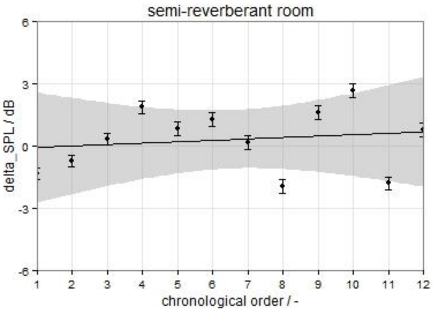


Figure 3 *ΔSPL by chronological task order with best fit linear regression lines in the semi-reverberant room.*

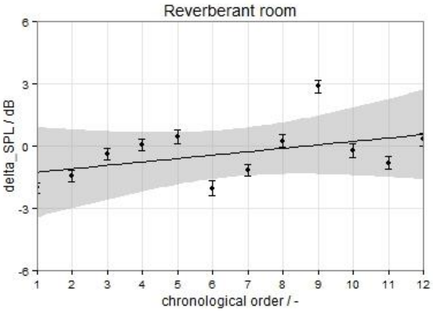


Figure 4 *ΔSPL by chronological task order with best fit linear regression lines in reverberant room.*

In order to adequately interpret these findings, it is necessary to consider the interaction between reverberation time and chronological task order. The relationship between SPL and apparent short-term vocal fatigue (evaluated by means of the chronological order of tasks from 1 to 12) was observed to strongly depend on room.

As reported in Equations (1), (2) and (3), the intercept for the anechoic room was -3.37dB (3.37dB lower than the typical vocal intensity), for the semi-reverberant room, 0.07dB, and for the reverberant room, -0.88dB. These values could be considered to be representative of the talker's impression of the room. In more unusual environments, such as anechoic and reverberant rooms, speakers could be intimidated by the environment and could therefore speak at a lower volume than usual. In the semi-reverberant room the intercept was very similar to the typical vocal intensity (0.07dB higher). As reported in Equations (1), (2) and (3), the slopes were 0.55 dB/Order in the anechoic room, 0.04 dB/Order in the semi-reverberant room and 0.08 dB/Order in the reverberant room. These values could be representative of the different effects of the room on short-term vocal fatigue; they suggest that lower vocal demands and a lower magnitude of vocal fatigue were experienced by talkers in the room in which the reverberation time was realistic, *i.e.*, the semi-reverberant room.

Estimates of the full effect of chronological task order over 12 tasks on ΔSPL, which were derived from the simplified models, were an increase of 6.05dB for the anechoic room, 0.44dB for the semi-reverberant room, and 0.88dB for the reverberant room. It could be predicted on the basis of these results that vocal effort in the reverberant room would have surpassed that in the semi-reverberant room if the subjects had been asked to speak for a longer period.

5. Conclusions

An experiment was conducted to measure the effect of reverberation time on short-term vocal fatigue. An analysis of vocal effort (SPL) indicated that speakers experienced less short-term fatigue in the presence of babble noise in the semi-reverberant room than in the anechoic and reverberant rooms.

6. Acknowledgements

The authors would like to thank the members of the Voice Biomechanics and Acoustics Laboratory, Michigan State University, for their support. Additionally, they would like to express their gratitude to the subjects involved in the experiment. This research was funded by the National Institute on Deafness and other Communication Disorders of the National Institutes of Health under 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.

7. References

[1] I. R. Titze, "Principles of voice production. Second Printing," *National Center for Voice and Speech, Iowa City*, pp. 229-233 & 361-366, 2000.

[2] D. Pelegrín-García and J. Brunskog, "Speakers' comfort and voice level variation in classrooms: Laboratory research," *J. Acoust. Soc. Am.*, vol. 132, pp. 249-260, 2012.

[3] E. J. Hunter, K. Tanner, and M. E. Smith, "Gender differences affecting vocal health of women in vocally demanding careers," *Logo. Phon. Voc.*, vol. 36, no. 3-4, pp. 128-136, 2011.

[4] É. Lombard, "Le signe de l'élévation de la voix," *Annales des Maladies de L'Oreille et du Larynx*, vol. 37, no. 2, pp. 101-119, 1911.

[5] ISO 9921: 2002(E), *Ergonomics – Assessment of speech communication*, International Organization for Standardization, Genève, 2002.

[6] H. Traunmüller and A. Eriksson, "Acoustic effects of variation in vocal effort by men, women and children," *J. Acoust. Soc. Am.*, vol. 107, no. 6, pp. 3438-3451, 2000.

[7] D. Pelegrín-García, B. Smits, J. Brunskog and C. Jeong, "Vocal effort with changing talker-to-listener distance in different acoustic environments," *J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 1981–1990, 2011.

[8] J. W. Black, "The effect of room characteristic upon vocal intensity and rate," *J. Acoust. Soc. Am.*, vol. 22, no. 2, pp. 174-176, 1949.

[9] A. Eriksson and H. Traunmüller, "Perception of vocal effort and speaker distance on the basis of vowel utterances," *Proc. of 14th ICPhS, San Francisco, California, USA*, 1999.

[10] L. Rantala, E. Vilkman, and R. Bloigu, "Voice changes during working: subjective complaints and objective measurements for female primary and secondary schoolteachers," *J. Voice*, vol. 16, no. 4, pp. 344-355, 2002.

[11] A. Laukkanen and E. Kankare, "Vocal loading-related changes in male teachers' voice investigated before and after a working day," *Folia Phoniatr. Log.*, vol. 58, pp. 229-239, 2006.

[12] J. Brunskog, G. Gade, G. Payà-Ballester, and L. Reig-Calbo, "Increase in voice level and speaker comfort in lecture rooms," *J. Acoust. Soc. Am.*, vol. 125, pp. 2072–2082, 2009.

[13] I. R. Titze, "Toward occupational safety criteria for vocalization," *Log. Phon. Vocol.*, vol. 24, pp. 49-54, 1999.

[14] E. J. Hunter and I. R. Titze, "Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings," *J. Sp. Lang. Hear. Res.*, vol. 53, pp. 862–875, 2010.

[15] P. Bottalico and A. Astolfi, "Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms," J. *Acoust. Soc. Am.*, vol. 131, no. 4, pp. 2817-2827, 2012.

[16] E. J. Hunter and I. R. Titze, "Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise," *Ann Otol Rhinol Laryngol.*, vol. 118, no. 6, pp. 449–460, 2009.

[17] E. Vilkman, E. R. Lauri, P. Alku, E. Sala, and M. Sihvo, "Effects of prolonged oral reading on F0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms," *J. Voice*, vol. 13, no. 2, pp. 303-312, 1999.

[18] M.P. Gelfer, M.L. Andrews, and C.P. Schmidt, "Effects of prolonged loud reading on selected measures of vocal function in trained and untrained singers," *J. Voice*, vol. 5, pp. 158, 1991.

[19] ISO 3382-2:2008(E), *Acoustics - Measurement of Room Acoustic Parameters, Part 2: Reverberation Time in Ordinary Rooms*, International Organization for Standardization, Genève, 2008.