Voices after midnight – How a night out affects voice quality

Alexandra Berger[1,2], Rosanna Hedström Lindenhäll[1,2], Mattias Heldner[2], Sofia Karlsson[1,2], Sarah Nyberg Pergament[1,2], Ivan Vojnovic[1,2],
[1] Department of Clinical Science, Intervention and Technology, Division of Speech and Language Pathology, Karolinska Institutet, Sweden
[2] Department of Linguistics, Stockholm University, Sweden
alexandra.berger@stud.ki.se, rosanna.lindenhall@stud.ki.se, mattias.heldner@ling.su.se, sofia.karlsson.3@stud.ki.se, sarah.nyberg.pergament@stud.ki.se, ivan.vojnovic@stud.ki.se

Abstract

This study aimed to investigate how different parameters of the voice (jitter, shimmer, LTAS and mean pitch) are affected by a late night out. Three recordings were made: one early evening before the night out, one after midnight, and one on the next day. Each recording consisted of a one minute reading and prolonged vowels. Five students took part in the experiment. Results varied among the participants, but some patterns were noticeable in all parameters. A trend towards increased mean pitch during the second recording was observed among four of the subjects. Somewhat unexpectedly, jitter and shimmer decreased between the first and second recordings and increased in the third one. Due to the lack of ethical testing, only a small number of participants were included. A larger sample is suggested for future research in order to generalize results.

Introduction

It is well known that the general volume at pubs, discotheques and similar venues is very loud and that as a guest you have to raise your voice significantly in order to make yourself heard. Speakers tend to raise their voices in loud conditions. This is known as the Lombard effect (Lane & Tranel, 1971). This type of voice behavior can result in vocal fatigue, temporary hoarseness and may in the long run cause vocal disorders (Vilkman, 2000).

This study aimed at examining how different acoustic voice quality parameters were affected by the voice strain induced by a night out talking in a noisy environment, and what effects can be observed the following day. The parameters examined in this study were jitter (cycle-to-cycle variations in frequency), shimmer (cycle-to-cycle variations in amplitude) (Titze, 1995), LTAS (long time average spectrum) and mean pitch.

Following Södersten, Ternström, & Bohman (2005) we expected the mean pitch to increase and that LTAS would indicate a decrease in vocal fry in the second recording. Furthermore, as previous results imply that female speakers tend to increase glottal closure after speaking in loud conditions (Linville, 1995), we hypothesized that jitter and shimmer would decrease continuously from the first to the third recording.

Method

To test our hypotheses we made three recordings and compared a number of voice quality measures in these. The first recording (R1) occurred at 7 pm on a Friday evening. The subjects each read a text of approximately one minute and then pronounced a prolonged [a]. The second recording (R2) took place at half past midnight, after four hours in a bar, where background noise level was measured. The third recording (R3) was done at noon the next day. The subjects reported differences in sleep duration (from 2 hours of sleep to 7 hours) as well as differences in alcohol intake.

Equipment

The recordings were done in 16-bit, 44.1 kHz with the application Røde rec LE (version 2.8.1) for iPhone 4 (version 7.0.3) and a Røde smartLav, tie clip, with a mouth-to-mic distance of 20 cm. Data was later analyzed in Praat (Boersma & Weenink, 2014). Noise level was measured with the application Buller (version 1.5) running on an iPhone 4S.

Subjects

The five participants consisted of four women and one man. Mean age was 26 years with standard deviation of 2.9 years. All of the subjects were speech and language pathology students from Karolinska Institutet. None of the five reported any voice problems. One of them, henceforth referred to as S1, smokes on a daily basis. All were informed of the potential health risks and participated voluntarily in the experiment.

Analysis

All voice quality analyses were performed in Praat (Boersma & Weenink, 2014). Mean pitch was measured for each one-minute text reading using the *To Pitch…* and *Get Mean…* functions in Praat. LTAS was calculated from the complete audio recordings (text reading plus vowels) in each session. The LTAS analyses were based on downsampled (10 or 11 kHz) and inverse filtered versions of the original audio recordings. The *To LPC (burg)* function in Praat was used for the inverse filtering. Perturbation measures of local jitter and shimmer were taken in the prolonged vowels, using the voice report function in Praat.

Differences in voice quality across the three recordings from each participant were tested using repeated measures ANOVAS. We used one-way repeated-measures ANOVAS to compare the effect of the recording session (R1, R2, R3) on three different voice quality measures: mean pitch, jitter, and shimmer. We used repeated contrasts to compare R1 vs. R2, and R2 vs. R3, respectively. Mauchly's test indicated that the assumption of spherizity was met in all three ANOVAS, therefore we will report the tests assuming spherizity below.

Results

Environmental noise

Measurements of the environmental noise were done repeatedly during one hour. These measurements showed that the background noise level varied between 80 and 92 dB(A), which is a normal noise level at these types of venues, but is indeed a strenuous environment for dialogue.

Mean pitch

Figure 1 shows the mean pitch in the different recording sessions for the individual subjects.

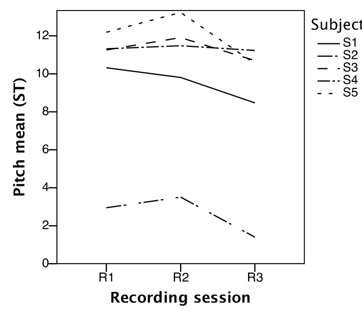

Figure 1. Mean pitch (in semitones relative to 100 Hz) in the three recording sessions for the individual subjects.

Evidently, four out of the five subjects had about 0.5 to 1 semitones higher pitch after midnight, and all subjects had a lower pitch on the day after although the amounts differed.

A one-way repeated-measures ANOVA showed that there was a significant effect of recording session on mean pitch (averaged across subjects), $F(2,8) = 10.41$, $p = .006$. Contrasts revealed that mean pitch was significantly lower in R3 than in R2, *F(1,4) = 13.36, p = .022,* and furthermore that R2 and R1 were not significantly different, *F(1,4) = 2.09, p = .222.*

LTAS

There was a lot of individual variation in the LTAS results for the five participants. Figure 2 shows an example from one subject (S5). For some of the subjects, there were clear differences between recordings while two of the participants showed little variation. Some subjects showed a more rapid decline within the first 1000 Hertz on R3 compared to the previous recordings indicating a steeper spectral slope.

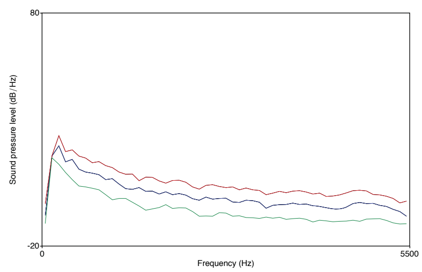

Figure 2. Example of LTAS curves from participant S5. The x-axis shows frequency (Hz) whilst the y-axis is showing sound pressure level (dB/Hz). The different lines represent the recordings: R1=middle line R2=upper line, R3=bottom line.

Jitter

Figure 3 shows the average jitter values in the different recording sessions for the individual subjects. All individual values were clearly below the *Multi-Dimensional Voice Program (MDVP)* jitter threshold of pathology of ≤1.040% (Kay Elemetrics, 2008). Somewhat unexpectedly, four out five participants had the highest jitter values in R1 and lower jitter value in R2 than in R1 and R3.

A one-way repeated-measures ANOVA showed that there was a significant effect of recording session also on jitter, *F(2,8) = 5.33, p = .034*. Contrasts revealed that jitter was significantly lower in R2 than in R1, *F(1,4) = 8.29, p = .045*, and furthermore that R2 also was significantly lower than R3, *F(1,4) = 15.30, p = .017.*

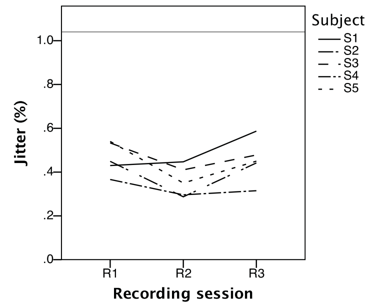

Figure 3. Jitter (in %) in the three recording sessions for the individual subjects. The grey horizontal line indicates the MDVP threshold of pathology for Jitter.

Shimmer

Figure 4 shows the average shimmer values in the different recording sessions for the individual subjects. All individual values but two were below the MDVP shimmer threshold of pathology of 3.810% (Kay Elemetrics, 2008). Again, unexpectedly, four out of five participants had the highest shimmer values in R1 and lower values in R2.

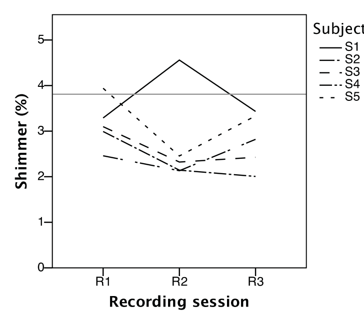

Figure 4. Shimmer (in %) in the three recordings sessions for the individual subjects. The grey horizontal line shows the MDVP threshold of pathology for Shimmer.

A one-way repeated-measures ANOVA showed that recording session did not have a significant effect on jitter,

$F(2,8) = 0.78$, $p = .49$. However, if the participant which behaved qualitatively different from the others was excluded, there was a significant effect, $F(2,6) = 5.60$, $p = .042$.

Discussion and conclusions

This study investigated the effects of speaking in a loud and noisy environment. Although, the results varied across subjects, certain recurring patterns were observed. As expected the mean pitch increased from R1 to R2 and decreased to R3. Surprisingly, all subjects except S1 decreased in both jitter and shimmer from R1 to R2 and increased to R3, although not to the same level as R1. Our theory is that the subjects were more vocally warmed up at R2, which might explain these results.

S1 differed from the others and increased in both jitter and shimmer during R2. This participant had results, which did not correlate with the others, even in pitch measures. We speculate that the individual differences can be explained by external factors such as alcohol consumption, cigarette smoking and amount of sleep. S1 had the largest intake of alcohol and cigarettes as well as only three hours of sleep.

Concerning LTAS there were not any strong differences between recordings, which may be due to environmental conditions during the recordings...

Since this is a pilot study with only a small number of participants it is difficult to get significant results. Also, because this study is explorative, it might be more interesting to look at the main effects of the experiment, rather than focusing on significance.

Because of the obvious problems in generalizing our results to a larger population we suggest a larger sample for future research. However, using a larger randomized sample might be hard to motivate ethically due to the possible health effects of this study.

We also suggest monitoring how the individual voices behave in loud environments in order to identify possible differences in voice behavior. Such differences might have an effect on the voice quality of the voices after a given occasion.

References

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] (Version 5.3.75). Retrieved from http://www.praat.org/

Kay Elemetrics. (2008). Multi-Dimensional Voice Program, Model 5105 [Computer program]. Lincoln Park, NJ, USA: Kay Elemetrics Corporation.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech, Language and Hearing Research*, *14*, 677–709.

Linville, S. E. (1995). Changes in glottal configuration in women after loud talking. *Journal of Voice, 9*(1), 57–65.

Södersten, M., Ternström, S., & Bohman, M. (2005). Loud speech in realistic enviromental noise: phonetogram data, perceptual voice quality, subjective ratings and gender differences in healthy speakers. *Journal of Voice, 19*(1), 29–46.

Titze, I. R. (1995). *Workshop on acoustic voice analysis: Summary statement.* Retrieved from http://www.ncvs.org/freebooks/summary-statement.pdf

Vilkman, E. (2000). Voice problems at work: A challenge for occupational safety and health arrangement. *Folia Phoniatrica Et Logopaedica, 52*(1-3), 120-125.