Silence and voicing accumulations in primary school teachers with and without voice disorders

Pasquale Bottalico[ac]

Simone Graetzer[a]

Arianna Astolfi[b]

Eric J. Hunter[a]

[a]Voice Biomechanics and Acoustics Laboratory

Department of Communicative Sciences and Disorders

Michigan State University

East Lansing, Michigan, United States of America

[b]Department of Energy

Politecnico di Torino

Torino, Italy

Running title: Silence and voicing accumulations

[c] Author to whom correspondence should be addressed. Electronic mail: pb@msu.edu

ABSTRACT

The relationship between vocal load, defined as the distribution of continuous silence and voicing periods, and subjects' clinical status was examined. Teachers were allocated by clinicians to groups: (1) with objectively measured vocal pathologies, (2) with subjectively/functionally reported symptoms but without objectively measured pathology, and (3) with normal physiology. Measurements were performed with the APM3200 during 4-hour workdays for 26 Italian primary school teachers. Silence and voicing accumulations were grouped into seven time intervals ranging from 0.03-0.9 s to 3.16-10 s according to Italian prosody. The greatest accumulations occurred in intervals ≥ 1.32 s for silence and in the middle intervals for voicing. Group 1 accumulated higher silence values in intervals between 0.1 and 3.15 s than other groups, while Groups 2 and 3 did not differ from each other. Silence accumulations < 3.16 s had no apparent effect on vocal recovery. Silence accumulations ≥ 3.16 s, which are necessary to ensure short term recovery (e.g., adequate fluid redistribution) in vocal fold tissue, were lower for pathological subjects. Voicing accumulations between 0.17 and 3.15 s were higher for pathological subjects. These results contribute to the understanding of the connection between voice disorders and vocal behavior in occupational voice users.

PACS Number: 43.70.Dn, 43.55.Hy, 43.70.Jt, 43.70.Mn, 43.72.Ar, 43.72.Dv.

I. INTRODUCTION

Voice disorders can be defined as conditions involving a variety of pathological symptoms that range from a mild disturbance of voice quality to complete loss of the ability to produce a laryngeal voice (Hillman, 2004). Such disorders are regularly experienced by occupational voice users because of the demands placed on the voice. Voice overuse is known to cause physiological vocal fatigue (Welham and Maclagan, 2003). Hunter and Titze (2009) state that laryngeal muscle fatigue results in soreness, discomfort, and/or muscle tension in the neck region, while tissue fatigue appears to be caused by change or damage to the vocal fold lamina propria following vibration exposure. Laryngeal tissue fatigue is associated with symptoms of pain or a scratchy voice sensation and/or increased voice breaks, instability and inability to produce soft voice.

Teachers comprise one of the occupational categories most affected by voice disorders (Titze *et al.*, 1997; Comins, 2002; Roy *et al.*, 2004; Kooijman *et al.*, 2006; Sliwinska-Kowalska *et al.*, 2006). While many studies have focused on occupational voice users in the United States, there are several studies examining the vocal behavior of non-English speaking teachers, specifically speakers of the Romance languages. For example, Angelillo *et al.* (2009) found that 60.1% of 504 Italian teachers reported suffering from voice problems. With regard to objectively identified vocal pathologies, studies employing laryngoscopic examinations have reported high rates of prevalence in teachers: 9.7% in Brazil (Filho *et al.*; 1995) and 13% in Spain (Urrutikoetxea *et al.*, 1995). Despite the prevalence of these problems, the occupational health and safety protocols for individuals in these professions are poorly developed (Villkman, 2000).

The vocal load of teachers has been characterized by several time dose studies, where time dose (*Dt*) refers to the time the vocal folds spend vibrating. For example, comparing the at-work vs not-at-work *Dt*, Hunter and Titze (2010) found that teachers vocalize on average for 30% of a 6 h work period, compared to 14.5% of a 6 h not at work period. Masuda *et al.* (1993) measured a mean phonation time of 21.6 % for elementary teachers and of 22.1 % for patients with vocal fold nodules and of 6.9 % for office workers over 8 h of work, while Bottalico and Astolfi (2012) found a mean voicing time percentage of 26% for primary school female teachers over a 4 h work period.

Hunter and Titze (2009) characterized a complete long-term recovery time on the basis of perceptual ratings on a 12 to 18 h period after a 2 h oral reading. They hypothesized that daily voice use leads to continual damage to the laryngeal tissue, and that the healing mechanism is in a state of constant repair. Hence, recovery time has a trajectory similar to that of a dermal wound healing trajectory (Robson *et al.*, 2001). As far as short recovery time is concerned, the minimum silence period for tissues to experience any degree of recovery has not yet been established.

In their study of vocal load and recovery, Titze *et al.* (2007) investigated the distributions of silence and voicing periods for teachers over the course of the day using an accelerometer, which was placed at the base of the subject's neck (Švec *et al.*, 2003; Popolo *et al.*, 2004). From the data collected during the work day, the average values of the occurrences and accumulations of silence and voicing periods per hour were obtained. The occurrences and the accumulations of silence and voicing periods were grouped into bin durations of half a decade of logarithmic time, according to English prosodic units, in the 0.0316 s to 31.6 s range for voicing and up to 103 s for silence. Over these frames, the durations of continuous silence and voicing periods were calculated. The resulting duration values, or silence and voicing periods, were assigned to logarithmic bins grouped into half decades. The occurrence of silence and voicing was counted per period and assigned to bins. The accumulation of silence and voicing was calculated as the product of the occurrence and the periods, and once again assigned to bins. The results showed that the greatest accumulation of voicing periods at work occurred in the (0.316–1.0) s range, and the greatest accumulation of silence, in the (3–10) s range. They argued that a minimum rest period of only a few seconds or a few minutes may be required, if an increase in blood circulation (Švec and Sram, 2001) or a redistribution of internal tissue fluid (Fisher *et al.*, 2001) occurs during this period of rest. Titze *et al.* (2007) did not consider in their study the clinical status of the subjects.

In the present study, silence and voicing accumulations at work of primary school teachers were related to the clinical status of the subjects. The primary aim was to determine whether and to what extent vocal pathologies affect teachers' vocal behavior, in particular, silence and voicing accumulations, during the work day. Previous research suggests a relationship between vocal pathology and improper breathing and abusive vocal behavior (Sapienza and Hoffman-Ruddy, 2009). Subjects with vocal fold nodules were found to have shorter periods of inhalation than healthy subjects (Iwarsson and Sundberg, 1999) and to have long phonation times (Masuda *et al*., 1993).

It was predicted in the present study that, during the workday, teachers with vocal pathologies would show (1) higher silence accumulations in shorter bins and lower silence accumulations in longer bins than teachers without pathologies, and (2) higher voicing accumulations than teachers without pathologies.

II. EXPERIMENTAL METHOD

The case studies concern 26 teachers at 7 primary schools in Italy: 14 teachers in 4 schools in Turin, which were built at the end of the nineteenth century, and 12 teachers in 3 schools in Beinasco, which were built in the 1970s. The subjects undertook voluntarily both the monitoring during lessons and the medical examinations. All subjects were native Italian speakers and traditional teachers (who teach classes of 20 to 30 pupils), with a mean age of 44.7 years (range 31-59). Teachers were monitored over 1 or 2 workdays of 4 h per day. The pupils' ages ranged between 6 and 11 years. A total of 43 workday samples were collected and all samples were included in the analyses. Table I reports the gender and age of the teachers and the number of monitored workdays. Special aid teachers were excluded from the study because their vocal load per day is substantially different from traditional teachers.

The acoustic conditions in the classrooms covered a wide range of reverberation time; the average values of mid-frequency reverberation time ranged between 0.6 s to 1.5 s in occupied conditions. The average background noise level, which did not differ significantly among the classrooms, was 50.6 dB(A). Acoustic conditions in the classrooms during phonation are reported by Bottalico and Astolfi (2012).

A. Clinical examinations

The teachers underwent clinical examinations, which were performed by a team of logopedists and phoniatricians, as described by Astolfi *et al.* (2012) and Vallino (2011). The examinations consisted of (1) a self-evaluation using the Voice Handicap Index (VHI-10, Jacobson *et al.*, 1997, Rose *et al.*, 2004), (2) a medical history (anamnesis), (3) an objective logopedic evaluation, and (4) a vocal health examination, which included phoniatric examinations and videolaryngostroboscopy (VLS). The medical history was obtained following the indications of Accordi and Tesserin (2002), while the objective logopedic evaluation was obtained following the indications of Vernero *et al.* (2002).

After medical examination, the logopedists and phoniatricians, who had clinical expertise (particularly with respect to diagnosis), evaluated the severity of the disorder, or the likelihood of the subject developing a disorder. Subsequently, the team of clinicians assigned subjects to the following groups: (1) 4 subjects (15.4%) with vocal pathologies detected both subjectively and objectively, with indications for therapy and speech treatment; (2) 11 subjects (42.3%) with either subjectively or functionally reported symptoms, but without objectively measured disease, with indications for vocal hygiene information and preventative speech treatment; and (3) 11 subjects (42.3%) with no pathological symptoms. Approximately 42% of the examined subjects showed no sign of disease, while 58% presented with subjectively and/or objectively measured pathological symptoms. These proportion are similar to those reported by Angelillo *et al.* (2009).

During the evaluation, each subject completed a VHI-10 assessment. The VHI-10 is a Likert scale of which each item is scored from 0 (*never*) to 4 (*always*), for a minimum of 0 and a maximum of 40; the higher the score, the more severe the patient's perception of disability due to a voice problem. In particular, a VHI-10 higher or equal to 11 should be considered abnormal (Arffa *et* al., 2012). Scores ranged between 0 and 19 of a possible 40. 19.2% of the subjects scored an abnormal value (higher or equal to 11). The mean value of the VHI-10 was 5.7 (s.d. 5.3).

As far as the objective evaluation of the vocal folds and larynx by means of VLS was concerned, 15 subjects were normally functioning ("normal physiology"), 4 presented with fold hypercontraction, 2 presented with hyperemia, 1, hypotonia, and 4 subjects presented with nodules and/or cysts.

Subjects were asked to report whether they had a hearing disorder. However, no hearing tests were performed. Subjects were therefore representative of the general teaching population.

In Table I, the self-reported hearing status, the VHI-10 score, the results of the objective evaluation of the vocal folds and larynx by means of VLS, and the subdivision into groups proposed by the team of clinicians is reported per subject.[1]

B. Measurements of silence and voicing accumulations

Each teacher was supplied with the Ambulatory Phonation Monitor (APM, model 3200, KayPENTAX®, Montvale, NJ). This device consists of an accelerometer, which was positioned below the talker's glottis at the sternal notch, and an acquisition unit that processed the accelerometer signal. The APM 3200 provided a time-history with a frame length of 50 ms. This time-history comprised the fundamental frequency, f_o, and an estimation of the sound pressure level, *SPL*, at a distance of 15 cm on-axis from the speaker's mouth, obtained after a calibration. The calibration was carried out by means of a reference microphone in order to correlate the skin acceleration level to the *SPL*.

Of the information provided by the device, only the detection of the presence or absence of voice excitation is of interest for the present study. Voiced and unvoiced frames were discriminated by the APM. When the RMS level acquired by the transducer exceeded a preset threshold, the frame was designated as voiced, and for that frame, f_o and *SPL* were determined (Cheyne *et al.*, 2003). Otherwise, the output result was equal to 0. The level acquired by the transducer was not affected by environmental noise. Silence and voicing accumulations, as defined in Sec. I, were derived from the time-histories provided by the APM.

[1] One of the subjects in the third group presented with a form of hypercontraction; however, according to the clinicians, it did not affect phonation.

The occurrences of continuous silence and voicing periods from 0.05 s to 10 s with a step of 50 ms were obtained from APM time histories. Subsequently, the accumulations for each time step were calculated by multiplying the occurrences by the corresponding step duration. The accumulation values were grouped into bins according to Italian prosodic units (Giordano, 2006; Romano, 2007; C-ORAL-ROM, 2005), as reported in Table II. Seven bins were used for the accumulations as follows: (Bin 1) 0.03-0.9 s long (silence and voicing periods below and up to the phonemic or segmental level); (Bin 2) 0.1-0.16 s long (at the level of unstressed syllables); (Bin 3) 0.17-0.33 s long (at the level of stressed syllables); (Bin 4) 0.34-0.66 s long (s at the word level); (Bin 5) 0.67-1.31 s long (at the non-terminal unit level); (Bin 6) 1.32 -3.15 s long (at the short tone unit level); (Bin 7) 3.16 - 10 s long (at the long tone unit level).

In addition, in order to better compare the results of the current study with Titze *et al.* (2007), a secondary analysis of the data was conducted in which the silence and voicing accumulations were allocated to bins in agreement with the bin widths specified by Titze *et al*. Specifically, the 6 bins used were as follows: (1) silence and voicing periods below and up to the phonemic segmental level (0.0316-0.10) s; (2) silence and voicing periods at the phonemic and syllabic level (0.10-0.316) s; (3) silence and voicing periods at the word and sentence level (0.316-1.0) s; (4) all-voiced sentences and pauses between sentences (1.0-3.16) s; (5) sustained phonations and pauses between sentences (3.16-10) s; (6) rare long phonations and silences in a dialogue (10-31.6) s. The longest bins were not considered because there were no accumulations in those bins by the subjects of the present study.

C. Statistical procedures

In order to characterize the dependence of silence and voicing accumulations on various covariates, linear mixed-effects models were fitted to the data. Such models can be said to have the form

$$Y_i = X_i\beta + Z_ib_i + \varepsilon_i \quad (1)$$

where Y_i represents a vector of responses for the *i*th group, X_i represents a fixed effects model matrix for group i, β represents a vector of fixed effects parameters, Z_i represents a random effects model matrix for group i, b_i represents a vector of random effects for group i, and ε_i represents a vector of errors. At least one of the random effects in the model represents the experimental units of the study, e.g., human subjects.

The model output includes the estimates of the fixed effects coefficients, β, the Standard Error associated with the estimate, the degrees of freedom (*df*), the test statistic, *t*, and the *p* value. The Satterthwaite method is used to approximate degrees of freedom and calculate *p* values. Typically, the parameters are estimated as those that minimize the restricted (or residual) maximum likelihood (REML) criterion. Information-theoretic metrics (including the Akaike information criterion) and the likelihood ratio test (LRT) are used to compare nested models and, in particular, to identify the most important predictors to be included in the models. Random effects terms are chosen on the basis of variance explained. Tukey’s post-hoc pair-wise comparisons are performed to examine the differences between all levels of the fixed factors of interest.

In this study, models were built and post-hoc comparisons were conducted using lme4, lmerTest and multcomp packages in R version 3.1.2 (R Development Core Team, 2011). Linear mixed models were chosen over linear models with log-transformed response variables and

Gamma models (with a log link) on the basis of (1) the distribution of points in residual vs. fitted value plots and (2) adjusted R^2 (proportion of variance explained). α was set at 0.05.

Kruskal-Wallis rank-sum tests (Kruskal and Wallis, 1952) were run in R 3.1.2 (R Development Core Team, 2011). This method is used to test for differences between the distributions of the observations (specifically the ranks of the observations) for two or more groups, without assuming normality of distribution. Between group sums of squares (representing between-group variance) are calculated from the average ranks. The test statistic, H, and the p-values, are approximated on the basis of a chi-square distribution. The null hypothesis is that the location parameters of the distributions are the same in each sample. The Benjamini-Hochberg (Benjamini and Hochberg, 1995) procedure can be used to control the false discovery rate.

The concept of Normalized Error (ISO/IEC Guide 43-1, 1997) was adopted for the analysis of compatibility between two sets of data, those reported in the present study, and those reported by Titze *et al*. (2007), which were obtained in different conditions, where no value(s) could be taken as the reference value(s). This test is used to determine whether the difference in the compared models is due to an effective difference between the evaluated phenomena or to systematic effects, rather than to random effects. The Normalized Error, E_N, is calculated as the ratio between the absolute value of the difference between the two samples mean and the relative expanded uncertainty of the difference (JCGM100, 2008), according to the following formula:

$$E_N = \frac{|m_1 - m_2|}{\mathrm{U}} = \frac{|m_1 - m_2|}{\mathrm{k}\sqrt{s_1^2 + s_2^2}} \qquad (2)$$

where m_1 and m_2 represent the average values of the two samples, s_1 and s_2 represent the standard deviations of the two samples and k is the coverage factor, calculated as the Student-t value for a conventional risk of error α of 5% and a number of degrees of freedom corresponding to $n-2$, where n is the number of samples used. This analysis can be considered a particular kind of hypothesis test. If the E_N value is higher than unity, the difference between the two sample means, m_1 and m_2, is higher than its uncertainty. Therefore, the difference is not merely due to random effects and the two results can be considered incompatible. Alternatively, if E_N is lower than unity, the difference could be due to random effects and there is no reason to reject the hypothesis of compatibility. Values lower than unity do not mean that real differences or systematic effects are not present, but rather that random effects cover their presence.

III. RESULTS AND DISCUSSION

A. Silence and voicing accumulations during the workday

Figure 1 shows the average values of silence and voicing accumulations in seconds per hour for each bin for the 26 subjects over the 43 workdays. Collectively, the data in Figure 1 represent 164.0 h of measurement. An average of 3.81 h per workday was measured, during which time the subjects were teaching pupils in a classroom. The average silence accumulation values were 47.0 s/h (Bin 1), 129.2 s/h (Bin 2), 62.5 s/h (Bin 3), 113.6 s/h (Bin 4), 184.0 s/h (Bin 5), 394.0 s/h (Bin 6) and 724.2 s/h (Bin 7). The peak of the silence distribution was in Bin 7 (3.16-10) s, which corresponds to silence periods at the long tone unit level. Average voicing accumulations were 30.9 s/h (Bin 1), 112.0 s/h (Bin 2), 182.4 s/h (Bin 3), 295.9 s/h (Bin 4), 162.9 s/h (Bin 5), 31.7 s/h (Bin 6) and 2.5 s/h (Bin 7). The greatest accumulation of voicing was found for Bin 4 (0.34-0.66) s, *i.e.*, the word level.

The results of the current study were compared to those of Titze *et al*. (2007). Figures 2 and 3 present a comparison of the silence and the voicing accumulation values obtained by Titze *et al.*

(2007) and the results obtained in the present study. Titze *et al.* found that the peak of the silence distribution was in Bins 5–6. The 3.16–31.6 s silence periods in Bins 5 and 6, which are typical of dialogue turn taking, were associated the greatest amount of accumulated vocal rest. The greatest accumulation of voicing at work time (451 s/h) was found for the word and sentence level, i.e., Bin 3 (0.316-1.0) s. In Figures 2 and 3 the accumulations obtained in the current study are reported in seconds per hour on a logarithmic scale and the bin widths are identical to those of Titze *et al.* (2007). In order to test the compatibility between the two studies, the Normalized Error values pertaining to the silence and voicing accumulations per Bin were calculated. All values were lower than one. In other words, the difference could be due to random effects and there is no reason to reject the hypothesis of compatibility. In sum, although their subjects taught a wider range of grades (from K to 12th grade), the results of Titze *et al.* (2007) are compatible with those of the present study.

B. Silence and voicing time percentages

On the basis of the accumulation distributions, the average percentage of the total monitoring time that the subjects spent in each bin was calculated for silence and voicing. With regard to silence periods, they spent 1.9% of the total time in Bin 1 (below and up to the phonemic segmental level), 5.2% in Bin 2 (unstressed syllable level), 2.5% in Bin 3 (stressed syllable level), 4.6% in Bin 4 (word level), 7.4% in Bin 5 (non-terminal unit level), 15.9% in Bin 6 (tone unit level) and 29.3% in Bin 7 (long tone unit level). With regard to voicing periods, they spent 1.2% of the total time in Bin 1, 4.5% in Bin 2, 7.4% in Bin 3, 12.0% in Bin 4, 6.6% in Bin 5, 1.3% in Bin 6 and 0.1% in Bin 7.

The trends in the accumulations are comparable with the findings for Swedish speakers of Löfqvist and Mandersson (1987). They measured a silence percentage of 15% for unvoiced segments, which is almost identical to the 14.2% measured in the present study (considering the first four bins, until the word level). Löfqvist and Mandersson found that in read monologues the voicing percentage was approximately 50%, and the silence percentage was 35% for boundary pauses. In the present analysis, the speech samples were not monologues but a mixture of monologue and dialogue. Consequently, the voicing percentage was lower (33.1%), with a higher percentage of boundary pauses (52.7%) due to the dialogue component. Boundary pauses were associated with Bins 5, 6 and 7, *i.e.*, periods longer than the word level.

The voicing percentage reported in the present study (33.1%) is similar to those obtained by Hunter and Titze (2010), in which 57 teachers were monitored over 2 weeks. They found that teachers vocalized for 29.9% of the occupational time, on average.

C. Effect of clinical status on accumulations

The effect of group on silence and voicing accumulations is shown in Figures 4 and 5, respectively. Two linear mixed-effects models were fitted for silence and voicing accumulations separately, each with 4 fixed and 2 random effects. The fixed effects were (1) Group, (2) Hearing condition, and interactions of (3) Bins and Group and (4) Bins and VHI-10 scores, divided into Normal and Abnormal levels. The models incorporated the following random effects structure: a random effect for Day (slope), indexed by Subject (intercept), so as to allow differing baseline levels of accumulations for subjects and differing responses by these subjects to the day of recording; and a random effect for Time step (intercept; in 50 ms steps; see Section IIB). Bins were treated as a continuous variable for the purposes of readily interpretable models. The results of the two models are shown in Table III. In both cases, no effect of hearing status was observed on accumulations. R^2 was 0.76 and 0.89 in the case of silence accumulations and voicing accumulations, respectively.

With regard to silence accumulations, the estimates of the differences between Group 1 and Groups 2 and 3 were $\beta = -16.2$ and $\beta = -17.5$, respectively, holding all other variables constant. The change in the slope silence accumulations-Bins between the normal and abnormal VHI-10 scores was not significant. The slopes for silence accumulations-Bins for Groups 1, 2 and 3 were $\beta = -6.7$, $\beta = -4.1$ and $\beta = -4.1$, respectively, where Group 1 was associated with higher values than Groups 2 and 3. Tukey's post-hoc multiple comparisons confirmed that pathological subjects (Group 1) accumulated higher values while Groups 2 and 3 behaved similarly (Group 2 – 1, $z = -17.8$, Group 3 – 1, $z = -18.6$, $p < 0.0001$; Group 3 – 2, $z = -1.3$, $p = 0.4$).

In the case of voicing, the difference in accumulations between the means for Group 1 and Groups 2 and 3 was $\beta = -25$ and $\beta = -26$, respectively, holding all other variables constant. The change in the slope voicing accumulations-Bins between the normal and abnormal VHI-10 scores was -0.28, reflecting higher values, especially in the central bins, for the high VHI-10 group. The relationship between voicing accumulations and the Voice Handicap Index (VHI-10) is shown in Figure 6. The slopes for voicing accumulations-Bins for Groups 1, 2 and 3 were $\beta = -10.3$, $\beta = -6.8$ and $\beta = -6.7$, respectively, where Group 1 was associated with higher values than Groups 2 and 3. Tukey's post-hoc comparisons confirmed that pathological subjects (Group 1) accumulated higher values while Group 2 and Group 3 behaved similarly to one another (Group 2 – 1, $z = -11.6$, $p<0.0001$; Group 3 – 1, $z = -12.5$, $p < 0.0001$; Group 3 – 2, $z = -1.6$, $p = 0.228$). The finding that VHI-10 scores predict voicing accumulations indicates a relationship between self-evaluated voice handicap and voice overuse.

Kruskal-Wallis tests (Table IV) with Benjamini-Hochberg adjusted *p* values identified differences between the Groups in Bins 2 to 7 for the silence accumulations, and Bins 3 to 6 for the voicing accumulations. In each case, with the exception of Bin 7 for the silence accumulations, there were higher accumulations for Group 1 than for Groups 2 and 3. In the case of Bin 7 of the silence accumulations (involving periods of ≥ 3.16 s), there were lower values for Group 1 than for Groups 2 and 3.

The finding of higher silence accumulations for pathological subjects (Group 1), with the exception of the longest bin, may relate to the respiratory and laryngeal functioning of pathological subjects. The respiratory behavior of pathological subjects can be considered from two points of view: (1) as a cause of pathology because a tendency towards shallow and quick breathing has been associated with vocal fold nodules (Iwarsson and Sundberg, 1999), apnea and muscle tension, and (2) as an effect of pathology because dysfunction in vocal fold adduction due to nodules (incomplete closure) can result in higher glottal airflow during phonation (Sapienza and Stathopoulos, 1994).

As mentioned, pathological subjects (Group 1) were associated with higher voicing accumulations, especially in bins 3 to 6. The overall accumulation of *Dt* was higher in Group 1 than in other groups (Group 1, 40.2%; Group 2, 31.9%; Group 3, 32.3%). It can be argued on the basis of these results that teachers with vocal pathology accumulate longer voicing periods than teachers without pathology. Vocal abuse is generally regarded to be the main cause of vocal fold nodules. Hence, as discussed in Section I, the vocal behavior of persons with long phonation times could be considered a factor in vocal abuse (Masuda *et al.*, 1993).

The present results indicate an increase of 54.9% in the voicing accumulations for Group 1 (subjects with objectively and subjectively measured vocal pathology) relative to Group 3 (normal physiology), and an increase of 4.1% for Group 2 (subjects with symptoms that were either subjectively or functionally identified, in the absence of an objectively measured disease) relative to Group 3.

IV. CONCLUSIONS

The primary aim of this study was to identify and characterize the differences among the distributions of silence and voicing periods during the workday for 26 primary school teachers with and without vocal pathology. Durations of continuous voicing and silence periods were calculated in 50 ms frames and assigned to bins grouped according to Italian prosodic units. Silence and voicing accumulations were calculated as a product of the occurrence (statistical frequency) and these periods. The teachers underwent a thorough clinical examination and were assigned to three groups: (1) subjects with vocal pathologies detected both subjectively and objectively; (2) subjects with either subjectively or functionally reported symptoms, but without objectively measured disease; and (3) subjects with no pathological symptoms. The effect of group on the silence and voicing accumulations was evaluated.

The main findings were as follows:

(1) Consistency was found between the results of the current study and those reported by Titze *et al.* (2007). In both studies, the highest peak of voicing occurred at 0.316 – 1 s (word and phrase boundary level, on their analysis) and of silence, at 3 – 10 s (pause between sentences, on their analysis).

(2) Subjects with objectively measured vocal pathology were associated with higher silence accumulations in the central bins, and lower silence accumulations in the longest bin, and higher voicing accumulations, especially in the central bins, than subjects without pathology, consistent with the predictions discussed in Section I.

(3) A relationship was observed between clinical status and *Dt*. Higher *Dt* was accumulated by subjects with objectively measured vocal pathology (40.6%) than other subjects (Group 2, 31.9%; Group 3, 32.3%).

(4) VHI-10 scores predicted voicing accumulations, such that subjects with abnormal scores were associated with higher voicing accumulations. There was no observable effect of VHI-10 on silence accumulations.

Titze *et al.* (2007) recognized that it is necessary to determine what rest period duration has a profound effect on vocal fatigue recovery. In the present study, as far as rest periods < 3.16 s were concerned, subjects with vocal pathologies displayed higher overall silence accumulations than others. The silence accumulations reported in this paper represent the accumulation of vocal rest during the workday. Hence, the results of this study indicate that rest periods shorter than 3.16 s may not have an observable effect on vocal fatigue recovery. With regard to rest periods ≥ 3.16 s, subjects with objectively measured vocal pathologies showed lower silence accumulations than other subjects. In this case, lower silence accumulations could indicate inadequate redistribution of fluids in the vocal fold tissue (Fisher *et al.*, 2001). It is feasible that this result may indicate an inadequate recovery time, which could lead to pathology.

The limitations of this paper include an imbalance in the sample sizes for the three groups, which was due to the voluntary nature of participation in the study and the fact that the clinical examination was conducted after subject selection and monitoring. Nevertheless, the proportions of subjects in the three groups in this study were likely to be representative of the proportions in the population (Filho *et al.*; 1995; Urritikoetxea *et al.*, 1995; Angelillo et al., 2009). In future work, clinical evaluation will be conducted prior to subject selection.

ACKNOWLEDGEMENTS

The kind cooperation of the teachers, children and school administrators has made this work possible. Thanks are due to L. Pavese for assistance in the initial stages of the analysis and to L. Hunter for editing assistance, and to T. Nesci, C. Marello, A. Romano and M. Moneglia for their scholarship in the area of Italian prosody. We are also grateful to have had the opportunity to collaborate with Clinic 1 ORL, Department of Surgical Sciences, University of Turin. Thanks are due especially to Doctors C. Giordano, A. Accornero, M. Garzaro, J. Nadalin, G. Pecorari, M. Garzaro, and L. Raimondo. In addition, the authors express their gratitude to those who provided comments and suggestions on an early version of this work, which was presented at the Fall 2014 meeting of the Acoustical Society of America. Analysis reported in this publication was partially supported 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.

REFERENCES

Accordi, M., and Tesserin, F. (**2002**). "L'anamnesi e l'esame obiettivo laringeo" ("Anamnesis and laryngeal objective examination"), in Proc. "Le disfonie: fisiopatologia, clinica ed aspetti medico-legali", LXXXIX Congresso Nazionale della Società Italiana di Otorinolaringologia e Chirurgia Cervico-Facciale (San Benedetto del Tronto, Italy).

Angelillo, M., Di Maio, G., Costa, G., Angelillo, N., and Barillari, U. (**2009**). "Prevalence of occupational voice disorders in teachers," J. Prev. Med. Hyg. **50**, 26-32.

Arffa, R. E., Krishna, P., Gartner-Schmidt, J., and Rosen, C. A. (**2012**). "Normative Values for the Voice Handicap Index-10," J. Voice **26**(4), 462-465.

Astolfi, A., Bottalico, P., Accornero, A., Garzaro, M., Nadalin, J., and Giordano, C. (**2012**). "Relationship between vocal doses and voice disorders on primary school teachers," in Proc. Euronoise 2012 (Prague, Czech Republic).

Benjamini, Y., and Hochberg, Y. (**1995**). "Controlling the false discovery rate: a practical and powerful approach to multiple testing," J. Royal Statistical Soc. Series B **57**, 289-300.

Bottalico, P., and Astolfi, A. (**2012**). "Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms," J. Acoust. Soc. Am. **131**(4), 2817-2827.

Comins, D. (**2002**). "Survey of UK voice clinics 2001/2," Voice Care Network UK.

Cheyne, H. A., Hanson, H. M., Genereux, R. P., Stevens, K. N., and Hillman, R. E. (**2003**). "Development and testing of a portable vocal accumulator," J. Speech Lang. Hear. Res. **46**(6), 1457-67.

C-ORAL-ROM (**2005**). *Integrated reference corpora for spoken romance languages*, edited by E. Cresti and M. Moneglia (John Benjamins Publishing Company, Amsterdam), pp. 1-304.

Filho, M., Gomez, F. G., and Macedo, C. (**1995**). “Videolaryngostroboscopy for pre-admissional examination of school teachers,” Proc. First World Congress of Voice, Oporto, Portugal.

Fisher, K. V., Telser, A., Phillips, J. E., and Yeates, D. B. (**2001**). “Regulation of vocal fold transepithelial water fluxes,” J. Appl. Physiol. 8750-7587 **91**(3), 1401–1411.

Giordano, R. (**2006**). “Note sulla fonetica del ritmo dell'italiano” (“Note on the Phonetics of rhythm in the Italian language”). In Proc. Analisi prosodica: Teoria, Modelli e Sistemi di annotazione. Atti del II Convegno Nazionale Associazione Italiana di Scienze della Voce (AISV) 2005, edited by R. Savy and C. Crocco (EDK Editore, Torriana), http://www.parlaritaliano.it/index.php/it/dati/450-note-sulla-fonetica-del-ritmo-dellitaliano (last viewed January 19, 2015).

Hillman, R. E. (**2004**). “Improving the assessment and treatment of voice disorders: emerging technologies,” Proc. From Sound to Sense, Cambridge, Massachusetts, USA.

Hunter, E. J., and Titze, I. R. (**2009**). “Quantifying vocal fatigue recovery: Dynamic vocal recovery trajectories after a vocal loading exercise,” Ann Otol Rhinol Laryngol. **118**(6), 449–460.

Hunter, E. J., and Titze, I. R. (**2010**). “Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings,” J. Speech Lang. Hear. Res. **53**, 862–875.

International Organization for Standardization (**1997**). ISO/IEC Guide 43-1 Proficiency Testing by Interlaboratory Comparisons. Part 1: Development and Operation of Proficiency Testing Schemes (International Organization for Standardization, Genève).

Iwarsson, J., and Sundberg, J. (**1999**). “Breathing behaviors during speech in healthy females and patients with vocal fold nodules,” Logoped. Phoniatr. Vocol. **24**(4), 154-169.

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., and Newman, C. W. (**1997**). "The Voice Handicap Index (VHI): Development and Validation," Am. J. Speech-Language Path. **6**(3), 66-70.

Joint Committee for Guides in Metrology (**2008**). JCGM100 Evaluation of measurement data—guide to the expression of uncertainty in measurement (Joint Committee for Guides in Metrology), http://www.bipm.org/en/publications/guides/gum.html (last viewed August 25, 2015).

Kooijman, P. G., De Jong, F. I., Thomas, G., Huinck, W., Donders, R., Graamans, K., and Schutte, H. K. (**2006**). "Risk factors for voice problems in teachers," Folia Phoniatr. Logop. **58**(3), 159-174.

Kruskal, W. H., and Wallis, W. A. (**1952**). "Use of Ranks in One-Criterion Variance Analysis," J. Amer. Statist. Assoc. **47**(260), 583-621.

Löfqvist, A., and Mandersson, B. (**1987**). "Long-time average spectrum of speech and voice analysis," Folia Phoniatr. (Basel) 0015-5705 **39**(5), 221–229.

Masuda, T., Ikeda, Y., Manako, H., and Komiyama, S. (**1993**). "Analysis of vocal abuse: fluctuations in phonation time and intensity in 4 groups of speakers," Acta Otolaryngol. **113**(4), 547-552.

Popolo, P. S., Švec, J. G., Hunter, E. J., Starr, A. C., Rogge-Miller, K., Nix, J., and Titze, I. R. (**2004**). "Teacher's guide to voice dosimetry," NCVS Online Technical Memo No. 4, October 2004: 1–31, http://www.ncvs.org/ncvs/library/tech/NCVSOnlineTechnicalMemo04.pdf (last viewed January 19, 2015).

R Development Core Team (**2011**). R: A language and environment for statistical computing, R Foundation for Statistical Computing (Vienna), http://www.R-project.org (last viewed January 19, 2014).

Robson, M. C., Steed, D. L., and Franz, M. G. (**2001**). “Wound healing: biologic features and approaches to maximize healing trajectories,” Curr. Probl. Surg. **38**(2), 72–140.

Romano, A. (**2007**). “À propos de deux bases de données de parole publiées récemment: compte-rendu de ‘API - Archivio del Parlato Italiano’ et de ‘C-ORAL-ROM’ (“On the matter of two recently published speech databases: a report on API and C-ORAL-ROM”), Rivista di Linguistica **19**(2), 367-390.

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T. and Murry, T. (**2004**). "Development and Validation of the Voice Handicap Index-10," Laryngoscope **114**(9), 1549-1556.

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., and Smith, E. M. (**2004**). “Prevalence of voice disorders in teachers and the general population,” J. Speech Lang. Hear. Res. **47**(2), 281-293.

Sapienza, C. M., and Hoffman-Ruddy, B. (**2009**). *Voice Disorders* (Plural Publishing, Evanston), p. 524.

Sapienza, C. M., and Stathopoulos, E. T. (**1994**). “Respiratory and laryngeal measures of children and women with bilateral vocal fold nodules”, J. Speech Lang. Hear. Res. **37**(6), 1229-1243.

Sliwinska-Kowalska, M., Niebudek-Bogusz, E., Fiszer, M., Los-Spychalska, T., Kotylo, P., Sznurowska-Przygocka, B., and Modrzewska, M. (**2006**). “The prevalence and risk factors for occupational voice disorders in teachers,” Folia Phoniatr. Logop. **58**(2), 85-101.

Švec, J. G., and Sram, F. (**2001**). "Violet-reddening of the vocal folds and the vocal fatigue," Proc. 25th World Congress Int. Assoc. Logoped. Phoniatr., Montreal, Canada.

Švec, J. G., Popolo, P. S., and Titze, I. R. (**2003**). "Measurement of vocal doses in speech: Experimental procedure and signal processing," Logoped. Phoniatr. Vocol. **28**(4), 181–192.

Titze, I., Lemke, J., and Montequin, D. (**1997**). "Populations in the U.S. Workforce Who Rely on Voice as a Primary Tool of Trade: A Preliminary Report," J. Voice **11**(3), 254-259.

Titze, I. R., Hunter, E. J., and Švec, J. G. (**2007**). "Voicing and silence periods in daily and weekly vocalizations of teachers," J. Acoust. Soc. Am. **121**(1), 469-478.

Urrutikoetxea, A., Ispizua, A., Mantellanes, F., and Aurrekoetxea, J. (**1995**). "Prevalence of vocal nodules in teachers," Proc. First World Congress Voice, Oporto, Portugal.

Vallino, D. (**2011**). "Clinical and acoustic analysis of vocal load of primary school teachers and speech therapy intervention," Thesis, Corso di Laurea in Logopedia, Facoltà di Medicina e Chirurgia, Università degli Studi di Torino, Italy.

Vernero, I., Gambino, M., Schindler, A. and Schindler, O. (**2002**). *La cartella logopedica - età adulta ed involutiva* (*Logopedic folder – Adulthood and Senescence*). (Omega, Turin).

Villkman, E. (**2000**). "Voice problems at work: A challenge for occupational safety and health arrangement," Folia Phoniatr. Logop. **52**(1-3), 120-125.

Welham, N. V., and Maclagan, M. A. (**2003**). "Vocal Fatigue: Current Knowledge and Future Directions," J. Voice **17**(1), 21-30.

FIGURE CAPTIONS

FIG. 1. Ensemble averages (over 42 workdays) of silence (in grey) and voicing (in white) accumulations per bin in seconds per hour with the standard error (SE) shown. The x-axis is bins corresponding to Italian prosodic units. The y-axis is accumulations in seconds per hour on a logarithmic scale.

FIG. 2. Barplots comparing the silence accumulation results of the present study (in white) with those (in grey) by Titze *et al.* (2007). The x-axis is accumulations assigned to logarithmic bins widths as specified by Titze *et al*. The y-axis is silence accumulations in seconds per hour on a logarithmic scale, with SD shown by error bars.

FIG 3. Barplots comparing the voicing accumulation results of the present study (in white) with those (in grey) by Titze *et al.* (2007). The x-axis is accumulations assigned to logarithmic bins widths as specified by Titze *et al*. The y-axis is silence accumulations in seconds per hour on a logarithmic scale, with SD shown by error bars.

FIG 4. Mean silence accumulations per group (Group 1 in grey, Group 2 in white, Group 3 in black) in seconds per hour with SD indicated by error bars. Group 1 consists of subjects with objectively measured vocal pathologies, Group 2, subjects with subjectively/functionally reported symptoms, and Group 3, without symptoms. The x-axis is bins corresponding to Italian prosodic units. The y-axis is silence accumulations in seconds per hour on a linear scale.

FIG. 5. Mean voicing accumulations per group (Group 1 in grey, Group 2 in white, Group 3 in black) in seconds per hour with SD indicated by error bars. Group 1 is the pathological group, group 2 are the subjects with reported symptoms, and group 3 is the healthy group. The x-axis is

bins corresponding to Italian prosodic units. The y-axis is voicing accumulations in seconds per hour on a linear scale.

FIG. 6. Voice Handicap Index (VHI-10) score by bin, where bins correspond to Italian prosodic units, with SD indicated by error bars. The x-axis is bins corresponding to Italian prosodic units. The y-axis is VHI-10 score.

Table I. Characteristics of the investigated teachers, Voice Handicap Index -10 scores, objective evaluation of the vocal folds and larynx by means of VLS and subdivision in groups proposed by a team speech pathologists and medical doctors.

Subject	*Gender*	*Age*	*Number of monitored workdays*	*Self-reported hearing condition*	*VHI-10 /40*	*VLS*	*Group*
1	Female	37	1	No	11	nodules	1
2	Female	34	1	No	9	nodules	1
3	Female	42	1	-	15	nodules and cysts	1
4	Female	54	1	Yes	19	bilateral nodules	1
5	Female	54	2	No	5	normal physiology	2
6	Female	49	1	Yes	5	hypercontraction	2
7	Male	59	2	No	14	vocal fold hyperemia	2
8	Male	43	2	No	1	normal physiology	2
9	Female	58	2	No	0	normal physiology	2
10	Female	33	2	No	5	normal physiology	2
11	Female	40	2	No	3	hypercontraction	2
12	Female	47	1	-	5	normal physiology	2
13	Female	54	2	Yes	7	vocal fold hyperemia	2
14	Female	43	2	No	0	hypotonia	2
15	Female	58	1	No	15	normal physiology	2
16	Female	34	2	No	3	normal physiology	3
17	Female	55	2	No	3	normal physiology	3
18	Female	52	2	No	3	normal physiology	3
19	Female	38	2	No	1	normal physiology	3
20	Female	56	2	Yes	2	normal physiology	3
21	Female	34	2	-	0	normal physiology	3
22	Female	39	2	No	5	normal physiology	3
23	Female	35	1	No	0	hypercontraction	3
24	Female	31	1	No	3	normal physiology	3
25	Female	40	2	Yes	5	hypercontraction	3
26	Female	38	2	Yes	10	normal physiology	3

Table II. Bin subdivision according to Italian prosodic units.

	Bin 1	*Bin 2*	*Bin 3*	*Bin 4*	*Bin 5*	*Bin 6*	*Bin 7*
Time Interval [s]	0.03-0.09	0.1-0.16	0.17-0.33	0.34-0.66	0.67-1.31	1.32 -3.15	3.16 -10
Voicing and Silence Period Correspondences	below and up to the phonemic segmental level speech	unstressed syllable level	stressed syllable level	word level	non-terminal unit level	short tone unit level	long tone unit level

Table III. Linear mixed models for response variables silence and voicing accumulations fitted by REML. The following fix factors are considered: (1) Group, (2) self-reported Hearing condition, and the interaction between (3) Group and Bins and (4) Bins and VHI-10.

Silence accumulations	*Estimate*	*Std. Error*	*df*	*t value*	*p value*	
(Intercept)	51.33	2.40	277.9	21.42	0.0001	***
Group2	-16.18	1.40	91.5	-11.57	0.0001	***
Group3	-17.48	1.40	92.6	-12.53	0.0001	***
Hearing condition	1.07	0.64	20.2	1.68	0.1076	
Bins:Group1	-6.65	0.35	309.1	-18.96	0.0001	***
Bins:Group2	-4.19	0.32	207.2	-13.23	0.0001	***
Bins:Group3	-4.10	0.32	206.0	-12.94	0.0001	***
Bins:VHI-10	-0.07	0.1	84.56	21.42	0.4709	

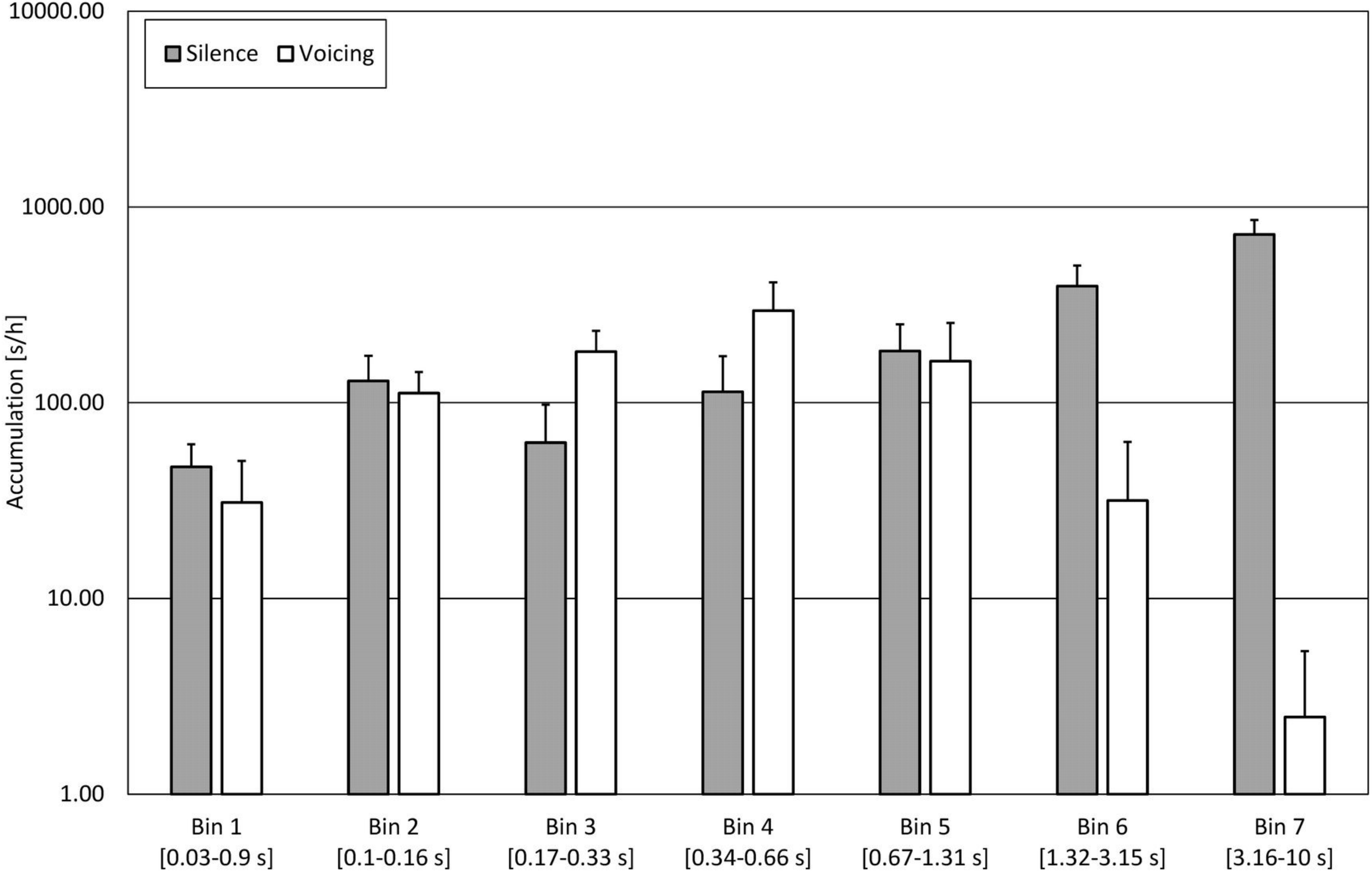
Voicing accumulations	*Estimate*	*Std. Error*	*df*	*t value*	*p value*	
(Intercept)	73.02	3.39	246.6	21.57	0.0001	***
Group2	-25.24	1.42	67.4	-17.80	0.0001	***
Group3	-26.25	1.41	68.0	-18.55	0.0001	***
Hearing condition	0.13	0.59	16.0	0.21	0.8327	
Bins:Group1	-10.30	0.51	240.4	-20.38	0.0001	***
Bins:Group2	-6.84	0.48	201.7	-14.14	0.0001	***
Bins:Group3	-6.74	0.48	201.3	-13.95	0.0001	***
Bins:VHI-10	-0.28	0.10	72.7	-2.81	0.0006	***

Signif. Codes: '***'<0.001 '**'<0.01 '*'<0.05

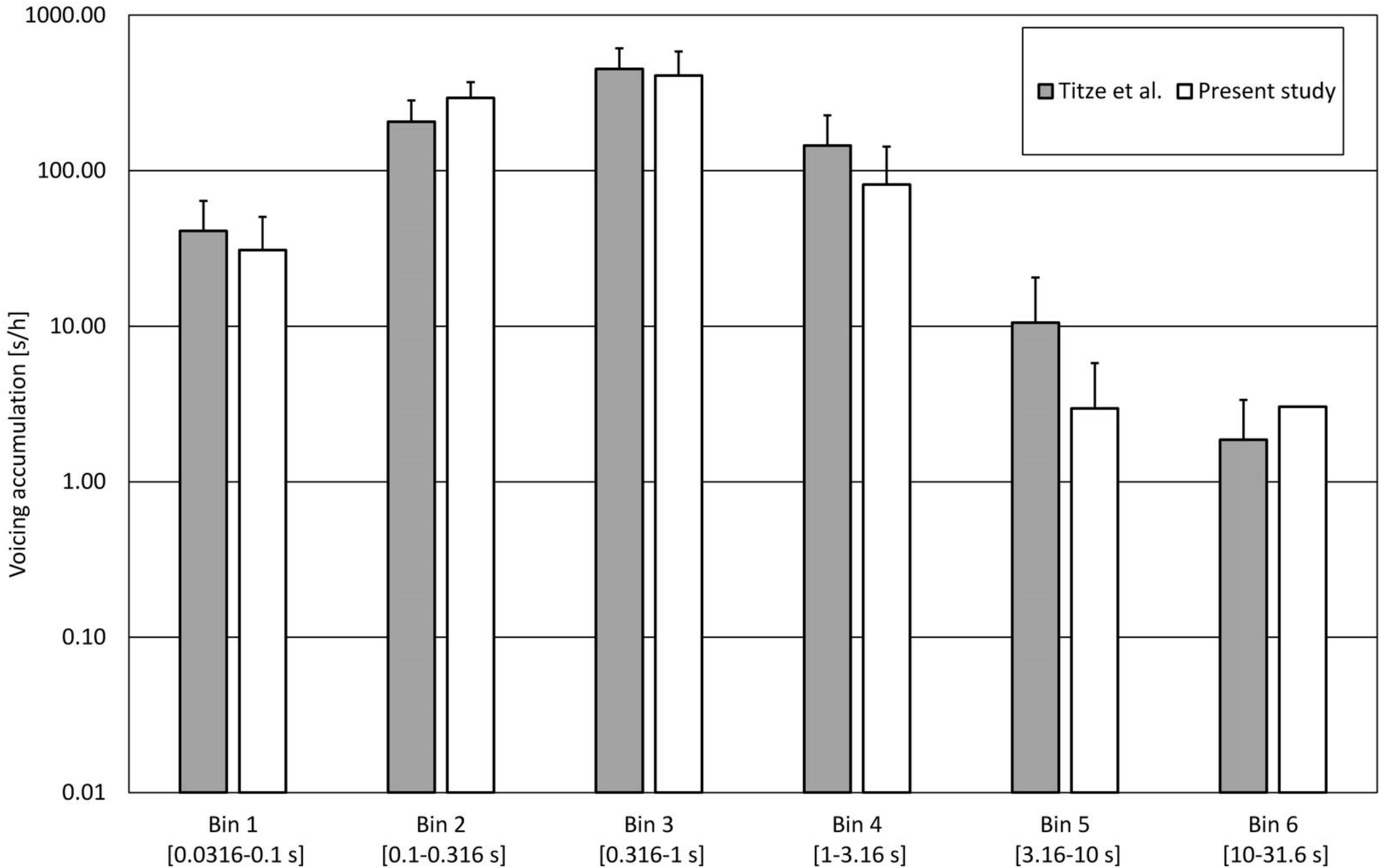
Table IV. Kruskal-Wallis test results (*H* test coefficient and Benjamini-Hochberg adjusted *p*-value) determining whether the location parameters of the silence and voicing accumulations were the same in each group. Degrees of freedom (*df*) = 2 in every case.

Silence accumulations				*Voicing accumulations*			
Bin	*H*	*p value*		*Bin*	H	*p value*	
1	3.728	0.155085		*1*	1.911	0.384557	
2	7.797	0.023646	*	*2*	6.074	0.067176	
3	9.18	0.014215	*	*3*	29.234	0.000001	***
4	31.071	0.000001	***	*4*	44.258	0.000001	***
5	46.916	0.000001	***	*5*	39.689	0.000001	***
6	53.897	0.000001	***	*6*	18.568	0.000163	***
7	31.556	0.000001	***	*7*	2.49	0.335922	

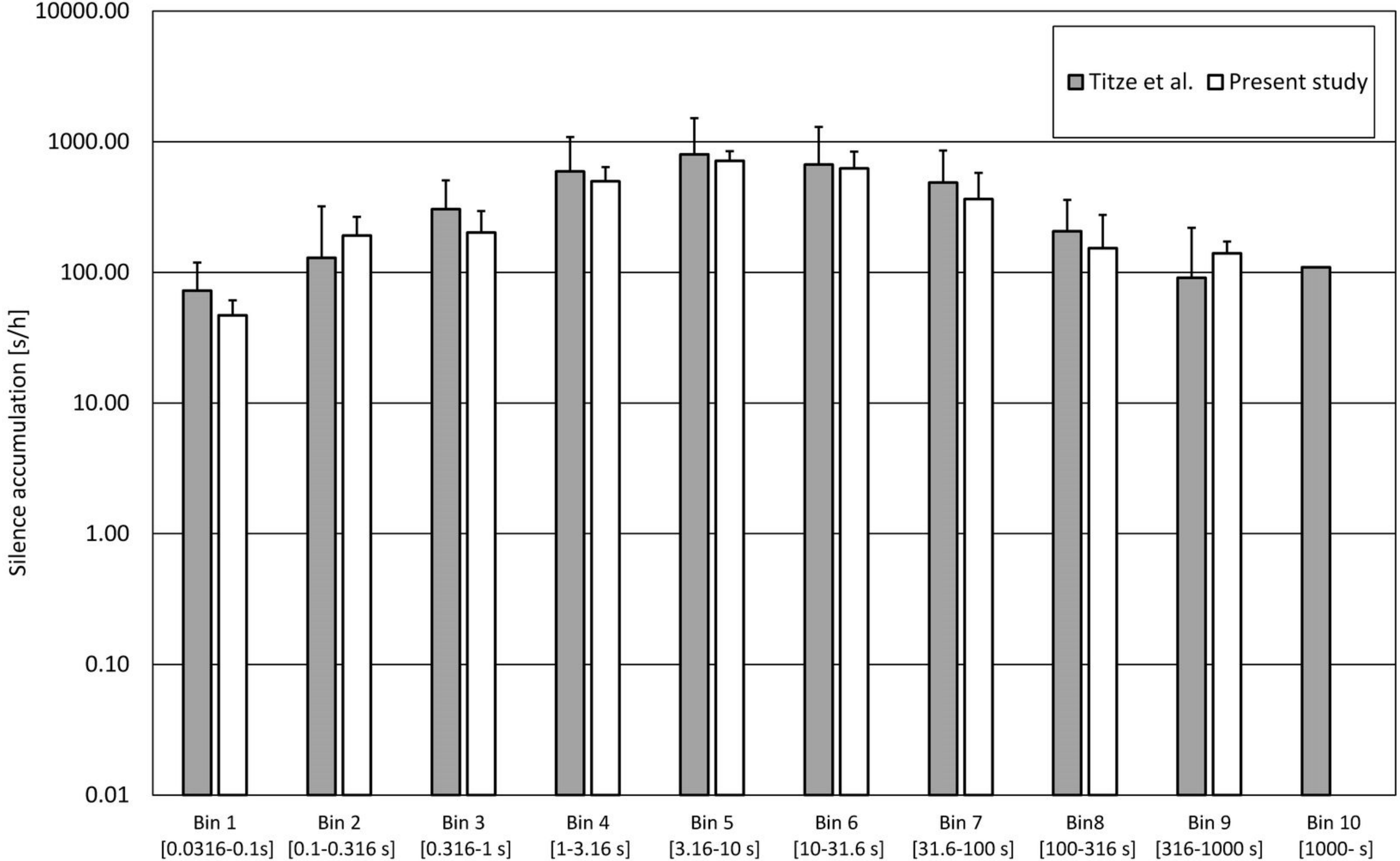
Signif. Codes: ’***’<0.001 ’**’<0.01 ’*’<0.05



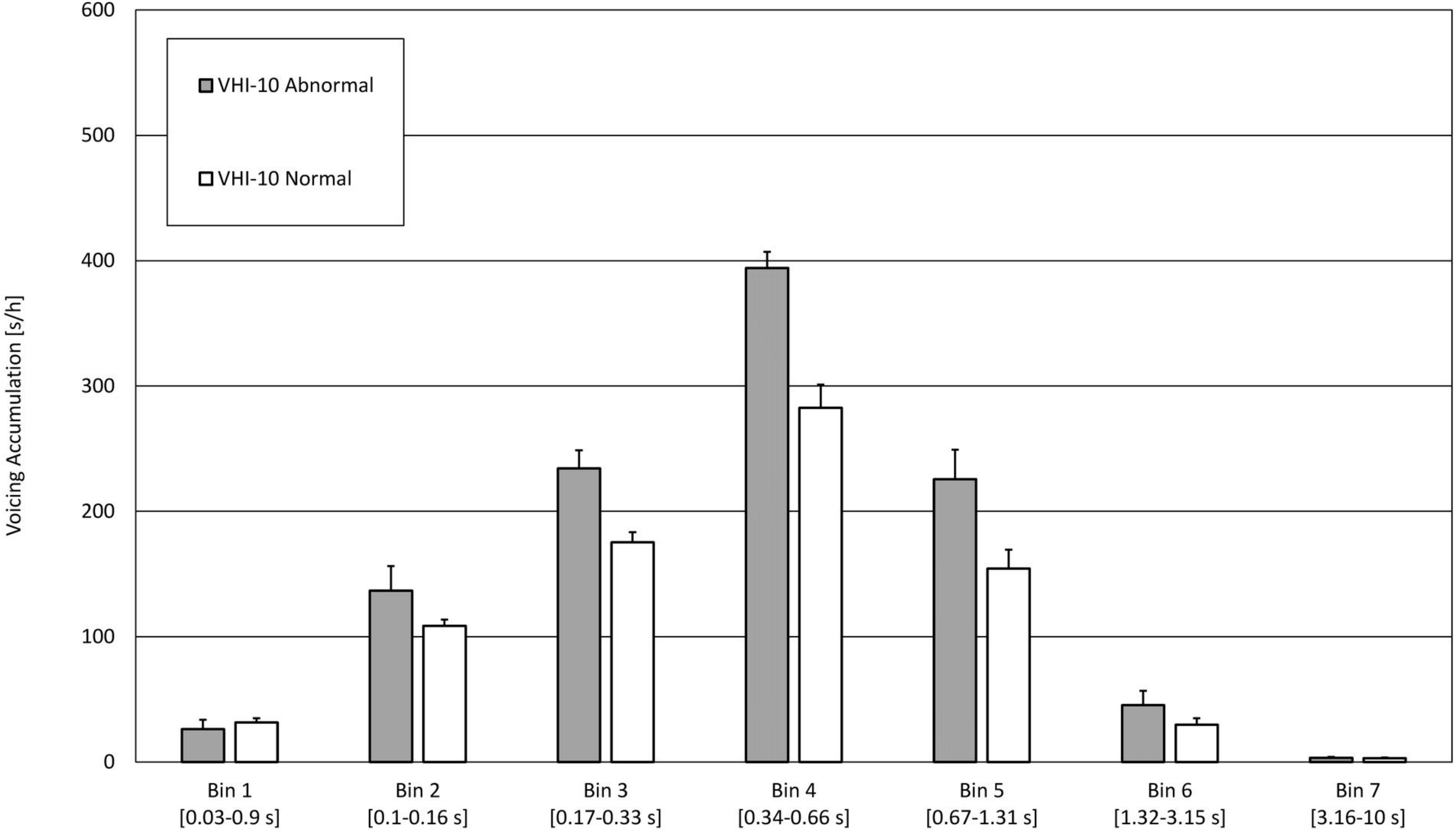
Silence
Voicing
10000.00
1000.00
100.00
10.00
1.00
Accumulation [s/h]
Bin 1
[0.03-0.9 s]
Bin 2
[0.1-0.16 s]
Bin 3
[0.17-0.33 s]
Bin 4
[0.34-0.66 s]
Bin 5
[0.67-1.31 s]
Bin 6
[1.32-3.15 s]
Bin 7
[3.16-10 s]

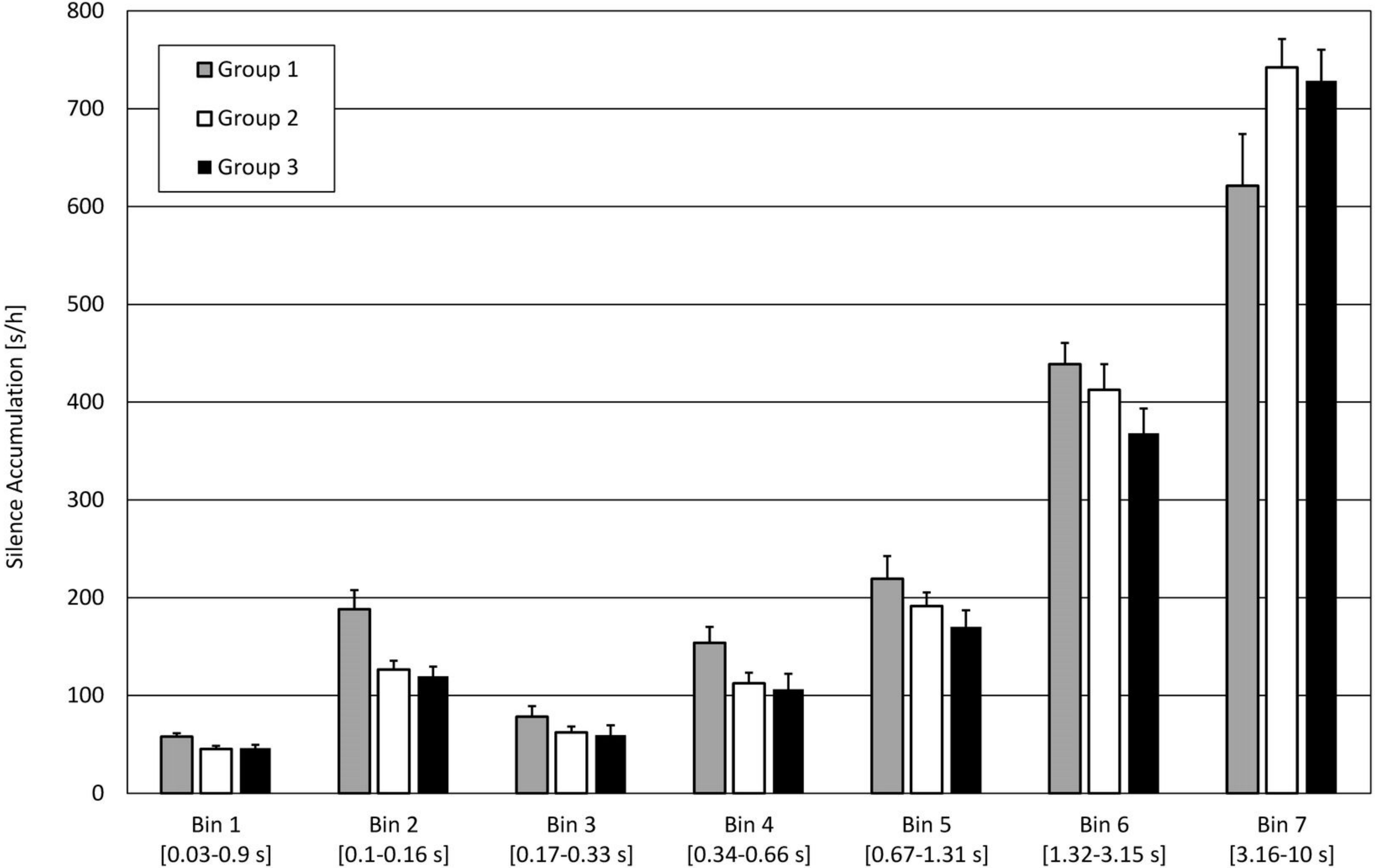


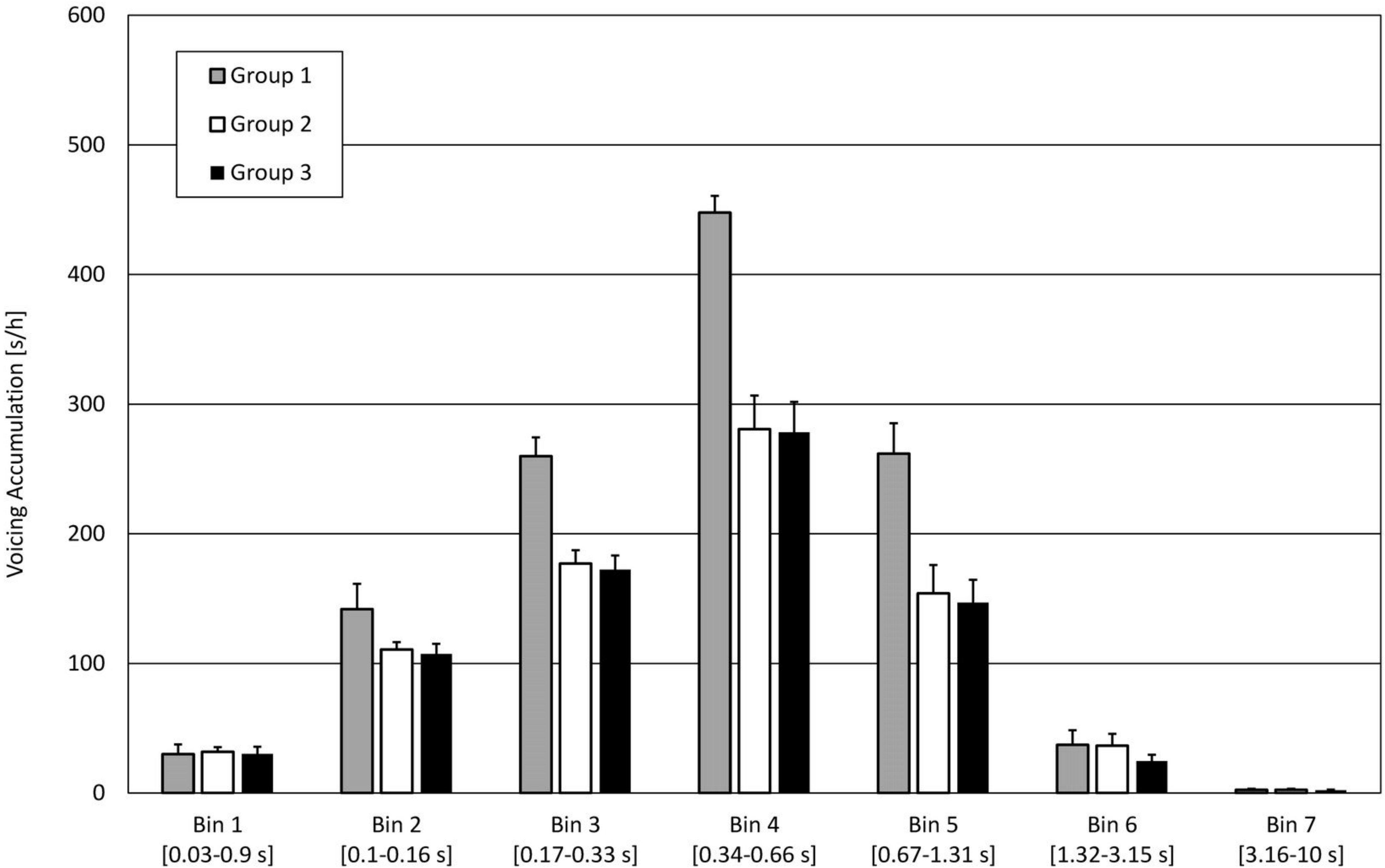
1000.00
100.00
10.00
1.00
0.10
0.01
Voicing accumulation [s/h]
Titze et al.
Present study
Bin 1
[0.0316-0.1 s]
Bin 2
[0.1-0.316 s]
Bin 3
[0.316-1 s]
Bin 4
[1-3.16 s]
Bin 5
[3.16-10 s]
Bin 6
[10-31.6 s]



Titze et al.
Present study
Silence accumulation [s/h]
10000.00
1000.00
100.00
10.00
1.00
0.10
0.01
Bin 1 [0.0316-0.1s]
Bin 2 [0.1-0.316 s]
Bin 3 [0.316-1 s]
Bin 4 [1-3.16 s]
Bin 5 [3.16-10 s]
Bin 6 [10-31.6 s]
Bin 7 [31.6-100 s]
Bin8 [100-316 s]
Bin 9 [316-1000 s]
Bin 10 [1000- s]


VHI-10 Abnormal
VHI-10 Normal
Voicing Accumulation [s/h]
600
500
400
300
200
100
0
Bin 1
[0.03-0.9 s]
Bin 2
[0.1-0.16 s]
Bin 3
[0.17-0.33 s]
Bin 4
[0.34-0.66 s]
Bin 5
[0.67-1.31 s]
Bin 6
[1.32-3.15 s]
Bin 7
[3.16-10 s]


Group 1
Group 2
Group 3
Silence Accumulation [s/h]
800
700
600
500
400
300
200
100
0
Bin 1
[0.03-0.9 s]
Bin 2
[0.1-0.16 s]
Bin 3
[0.17-0.33 s]
Bin 4
[0.34-0.66 s]
Bin 5
[0.67-1.31 s]
Bin 6
[1.32-3.15 s]
Bin 7
[3.16-10 s]



Voicing Accumulation [s/h]
600
500
400
300
200
100
0
Group 1
Group 2
Group 3
Bin 1
[0.03-0.9 s]
Bin 2
[0.1-0.16 s]
Bin 3
[0.17-0.33 s]
Bin 4
[0.34-0.66 s]
Bin 5
[0.67-1.31 s]
Bin 6
[1.32-3.15 s]
Bin 7
[3.16-10 s]