JASA ARTICLE

Modeling voice production and self-perception in noise: Understanding the Lombard effect in non-phonotraumatic vocal hyperfunction[a)]

Christian Castro,[1] Juan P. Cortés,[2] Lucía Z. Rivera,[2] Pavel Prado,[3] Alejandro Weinstein,[2,4] and Matías Zañartu[2,4,b)]

[1]*Rehabilitation Sciences Institute, School of Speech Therapy, Faculty of Rehabilitation Sciences, Universidad Andres Bello, Santiago, 7591538, Chile*

[2]*Advanced Center of Electrical and Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile*

[3]*Escuela de Fonoaudiología, Facultad de Odontología y Ciencias de la Rehabilitación, Universidad San Sebastián, Santiago 7510602, Chile*

[4]*Department of Electronic Engineering, Universidad Técnica Federico Santa María, Valparaíso, Chile*

ABSTRACT:

The sensorimotor adaptation process is crucial for maintaining oral communication. Recent studies have shown that individuals with non-phonotraumatic vocal hyperfunction (NPVH) experience difficulties in sensorimotor adaptation when speaking in noise (known as the Lombard effect). However, the role of auditory and somatosensory feedback in the dynamics of adaptation to speaking in noise is still unclear. In this study, the use of a simple three-parameter mathematical model, known as SimpleDIVA model, was extended to explore the adaptation dynamics of speaking in noise among a group of participants with typical voices and NPVH. All participants were asked to utter a series of syllables under three conditions: baseline (quiet environment), Lombard (speech-shaped noise at 80 dB), and recovery (quiet environment after 5 min of rest). The results indicate that participants with NPVH did not return to baseline after exposure to speaking under noise. The SimpleDIVA model analysis reveals a diminished feedforward learning rate and reduced somatosensory feedback gain in participants with NPVH in comparison to participants with typical voices. This suggests that participants with NPVH may be using less somatosensory information when speaking in noise and may require more time to update the feedforward commands during and after speaking in noise. https://doi.org/10.1121/10.0034544

(Received 20 February 2024; revised 3 November 2024; accepted 7 November 2024; published online 6 December 2024)

[Editor: Jody Kreiman] Pages: 3772–3779

I. INTRODUCTION

When individuals speak in noisy environments, they involuntarily increase their vocal intensity, which is a phenomenon known as the Lombard effect (LE; Lombard, 1911). This effect is driven by auditory feedback processes that help maintain a signal-to-noise ratio (SNR) sufficient for intelligibility in challenging acoustic conditions (Lane and Tranel, 1971; Pick *et al.*, 1989; Tartter *et al.*, 1993). However, the mechanisms underlying this phenomenon are not well understood in individuals with voice disorders, where different processes may be involved. One such disorder is non-phonotraumatic vocal hyperfunction (NPVH), which is characterized by difficulties in sensorimotor adaptation when speaking in noise, leading to distinct patterns compared to controls (Castro *et al.*, 2022b). In these individuals, the SNR may be a less effective acoustic cue, or they may rely on different speech motor control mechanisms. In this study, we employ a model-based approach to investigate how sensorimotor adaptation contributes to the LE in individuals with typical voices and those with NPVH.

Acoustic cues play a critical role in speech production and perception, particularly in challenging environments (Huilgol *et al.*, 2019; Kalikow *et al.*, 1977; Stevens, 2002). The auditory system relies on these cues to discriminate speech amidst noise, whereas speech motor control adjusts vocal output to ensure effective communication. This dynamic interplay between sensory information and motor adjustments is central to sensorimotor adaptation.

In this context, sensorimotor adaptation involves adjusting motor responses based on sensory feedback to maintain performance accuracy (Houde and Jordan, 1998). In speech production, stable auditory and somatosensory conditions lead to reduced reliance on feedback as the feedforward system becomes robust. However, prolonged changes in sensory conditions prompt updates to the feedforward system to correct discrepancies between predicted and actual feedback, as observed in the "aftereffect" in speech adaptation experiments (Behroozmand and Sangtian, 2018; Houde and Jordan, 1998; Houde and Nagarajan, 2011).

Vocal hyperfunction (VH) is an etiological component of several voice disorders and is characterized by excessive

[a)]This paper is part of a special issue on Acoustic Cue-Based Perception and Production of Speech by Humans and Machines.

[b)]Email: matias.zanartu@usm.cl

or unbalanced laryngeal muscle activity, leading to increased vocal effort and irregular vocal fold vibrations (Aronson and Bless, 2009; Hillman *et al.*, 2020). NPVH, a subtype of VH, involves heightened laryngeal muscle activity without vocal fold tissue trauma (Hillman *et al.*, 2020). NPVH is associated with factors such as inefficient phonatory function, psychological stress, impaired auditory discrimination, and sensorimotor deficits (Abur *et al.*, 2021; Demmink-Geertman and Dejonckere, 2002; McKenna *et al.*, 2020; Stepp *et al.*, 2017). Recent research has shown that individuals with NPVH struggle with sensorimotor adaptation when speaking in noise, exhibiting higher voice intensity even after noise cessation (Castro *et al.*, 2022b). This aftereffect suggests challenges in updating feedforward commands in response to varying background noise levels. However, the dynamics of sensorimotor adaptation in noise remain unclear, requiring advanced tools for investigation.

The DIVA model (directions into velocities of articulators) is a well-established framework for studying speech motor control (Cuadros *et al.*, 2023; Guenther *et al.*, 1998; Tourville and Guenther, 2011). It proposes a feedforward system that generates predictive motor commands and a feedback system that incorporates auditory and somatosensory input to correct discrepancies. The SimpleDIVA model, a simplified version of DIVA, focuses on sensorimotor adaptation using altered feedback paradigms and has been applied in studies of pitch and formant shifts (Kearney *et al.*, 2020). However, it has not yet been used to explore adaptation to speaking in noisy conditions.

This study aims to explore how speaking in noise (LE) influences sensorimotor adaptation by examining the interactions between auditory feedback, somatosensory feedback, and feedforward command updates in individuals with typical voices and those with NPVH using the SimpleDIVA model. Given the challenges that participants with NPVH face in adapting to noise, we hypothesize that they may not rely on SNR as the primary acoustic cue in response to increased background noise.

II. MATERIALS AND METHODS

A. Participants

Forty participants were recruited for this study: 20 volunteers with typical voice (control group) and 20 participants with NPVH. The mean (standard deviation) of participants' ages was 28 years old (3.2 years) for individuals with typical voices and 29 years old (2.8 years) for those with NPVH. All participants were assessed by a speech-language pathologist and otolaryngologist based on case history, clinical evaluation, laryngeal endoscopy, aerodynamic and acoustic measures of vocal function, and a Spanish version of Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Núñez-Batalla *et al.*, 2015). Moreover, all participants passed a pure-tone hearing screening, which consisted of measuring the auditory threshold to air-conduction stimuli in both ears at octave frequencies between 250 and 8000 Hz using a clinical audiometer (model AD629, Interacoustics A/S, Middelfart, Denmark). To pass the screening, each participant was required to have an auditory threshold below 20 dB hearing level across all relevant frequencies.

B. Experimental design

The experiment involved participants sitting in front of a screen and a microphone placed 15 cm in front of their lips. Each participant was then asked to utter 80 syllables at a comfortable pitch and loudness. The syllables included /pa/, /da/, /ta/, and /ba/ presented randomly. The duration of vocalization (3 s) and pace of syllable pronunciation (one syllable every 6 s) were controlled by visual cues displayed on a screen (Fig. 1). Participants were instructed to speak at a comfortable loudness level while receiving auditory feedback of their own voice through headphones. No specific loudness or pitch was targeted to appropriately assess the LE. The process was repeated under three sequential acoustic background conditions: baseline (in quiet), Lombard (in noise), and recovery (in quiet after 5 min of rest).

The noise used for the Lombard condition consisted of speech-shaped noise generated by a clinical audiometer (model AD629, Interacoustics A/S, Middelfart, Denmark) and presented through the headphones at a sound pressure level (SPL) of 80 dB (dB re 20 μPa) This noise is characterized by maintaining equal energy between 125 and 1000 Hz (octave band), followed by an energy decay of 12 dB/octave until 6000 Hz. The selection of this noise level is based on previous studies that demonstrated that it is effectiveness in triggering vocal adaptations in response to noisy environments (Castro *et al.*, 2022b; Garnier *et al.*, 2010; Junqua, 1993; Lu and Cooke, 2008; Summers *et al.*, 1988). This level is sufficiently loud to induce a robust LE while avoiding discomfort and vocal/auditory fatigue (Alghamdi *et al.*, 2018). Moreover, it is important to note that speakers are often exposed to higher levels of noise in their daily activities. In this context, previous studies have reported noise levels in the range of 60–85 dB in restaurants and eating establishments (Bottalico, 2018; Hodgson *et al.*, 2007; Nahid and Hodgson, 2011). Similarly, school classrooms have been shown to exhibit higher noise levels, reporting a range between 45 and 87 dB (Darius *et al.*, 2023; Lamotte *et al.*, 2021; Mealings *et al.*, 2024; Xia *et al.*, 2024). Additionally, noise levels at train stations have been reported to reach up to 90 dB (Younes *et al.*, 2021). These environments are generally characterized by many people speaking simultaneously at higher levels. For this reason, the type of noise used in this experiment contained speech frequencies and intensity levels similar to those encountered in real-life situations. The background noise in the quiet condition of the soundproof booth was 35 dB SPL. The recovery condition was included to explore the potential persistence of the LE after speaking under noise (Behroozmand and Sangtian, 2018; Castro *et al.*, 2022b). Previous studies have regarded speaking in noise as a vocal

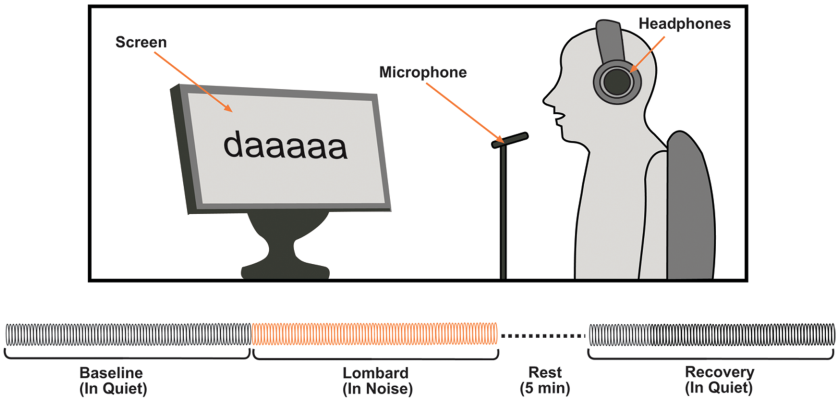

FIG. 1. (Color online) An overview of the experimental setup shows a person seated in front of a monitor. The individual is wearing headphones through which they hear their own voice in real-time. While performing the different stages of the experiment, a microphone records their voice. Below, we can see a representation of the experiment, illustrating the three conditions: baseline (in quiet), Lombard (in noise), and recovery (in quiet).

loading process (Fujiki and Sivasankar, 2017) and described the return to baseline for acoustic parameters as occurring within 5 min of vocal rest after noise removal in quiet conditions in participants with typical voices (Castro *et al.*, 2022b; Xue *et al.*, 2019). Therefore, the experiment considered a 5-min rest between the Lombard and recovery conditions to allow participants to recover from potential vocal loading effects, as suggested in previous studies (Castro *et al.*, 2022b; Fujiki and Sivasankar, 2017; Xue *et al.*, 2019).

The acoustic signal was obtained using a microphone (BK, model 4961; Naerum, Denmark) located in front of the participant at 15 cm from the lips at a 45-deg offset in the axial direction and amplified by a BK 1705 signal conditioner. The acoustic signal was calibrated to physical units of dB SPL using a Larson Davis calibrator (model CAL200, Depew, NY). Next, signals were sampled at 20 kHz with 16-bit quantization and low-pass filtered (3 dB cutoff frequency of 8 kHz) using a National Instruments DAQ model USB-6363 BNC (Austin, TX). From the calibrated acoustic signal, we estimate the SPL for each vocalization using a window of 200 ms from the stable part of each vocalization.

A two-way repeated measures analysis of variance (ANOVA) was conducted to examine the variation in difference-SPL across the experimental conditions. The factors in the ANOVA were condition, group, and the interaction between group and condition. A *post hoc* Tukey multiple comparison test was performed to assess the statistical significance between conditions ($p < 0.05$). Prior to this, tests for normality (Shapiro-Wilk test) and homogeneity of variances (Levene test) were conducted to verify the assumptions.

C. Extended SimpleDIVA model for speaking in noise

The SimpleDIVA model proposes three equations related to key subsystems involved in speech motor control: auditory feedback control, somatosensory feedback control, and feedforward control,

$$F_{\text{produced}}(n) = F_{FF}(n) + \Delta F_{FB}(n), \quad (1)$$

$$\Delta F_{FB}(n) = \alpha_A(F_T - F_{AF}(n)) + \alpha_S(F_T - F_{SF}(n)), \quad (2)$$

$$F_{FF}(n+1) = F_{FF}(n) + \lambda_{FF}\Delta F_{FB}(n). \quad (3)$$

The feature value $F(n)$ at trial n is a vectorial feature of D dimensions. Experiments reported with SimpleDIVA use a feature acoustic dimension (e.g., fundamental frequency, one or more formants, etc.), which generally has a percentage of perturbation from an initial target value (F_T). Usually, the average of F during the baseline phase, where the participant elicits the utterance without perturbation (see Sec. II for more details) for all experiments. Equation (1) states that the elicited $F(n)$ for a trial n is the sum of the feedforward $F_{FF}(n)$ plus a correction factor $\Delta F_{FB}(n)$. This correction factor is defined in Eq. (2), which is a weighted sum between the errors in auditory and somatosensory feedback control, where each one is multiplied by auditory and somatosensory feedback gains, α_A and α_S, respectively. Last, Eq. (3) is the update for the feedforward command to the next trial $F_{FF}(n+1)$, which is the sum of the current command $F_{FF}(n)$ plus a fraction of the correction factor from Eq. (2). The fraction of $\Delta F_{FB}(n)$ is controlled by the learning rate λ_{FF}. For more details on the development of these equations, please refer to Kearney *et al.* (2020). The parameters α_A, α_S, and λ_{FF} can be multidimensional and are optimized using particle swarm optimization (Kennedy and Eberhart, 1995) from the dataset. These optimized parameters represent the contribution of the auditory feedback, somatosensory feedback, and feedforward control, respectively, from the SimpleDIVA model.

Most adaptive experiments reported with SimpleDIVA use features related to the fundamental frequency (f_o) and/or the first and second formants (F_1, F_2; Kearney *et al.*, 2020). In contrast, the feature of interest in the present work is the SPL, specifically due to a change in background noise that reduced the SNR by 45 dB (from 35 dB in baseline to 80 dB in Lombard). In this context, our perturbation parameter in SimpleDIVA is defined as the ratio of the expected acoustic power in the Lombard condition to the observed power in the baseline condition

$$\text{SPL-pert} = \frac{P_T^2}{P_L^2} - 1, \quad (4)$$

where P_T represents the acoustic pressure derived from the average SPL during the baseline condition, and P_L is the expected acoustic pressure during the Lombard condition.

The latter assumes an incoherent sum, expressed as $P_L^2 = P_T^2 + P_N^2$, where P_N is the acoustic pressure of the noise (i.e., $\text{SPL}_N = 80$ dB), which the speakers are exposed to through the headphones. Given these definitions and the relationship between SPL and acoustic pressure, the SPL perturbation parameter can be further expressed as a function of SPL_T and SPL_N such that

$$\text{SPL-pert} = \frac{P_T^2}{P_T^2 + P_N^2} - 1 = -\frac{P_N^2}{P_T^2 + P_N^2} = -\frac{10^{\text{SPL}_N/10}}{10^{\text{SPL}_T/10} + 10^{\text{SPL}_N/10}}. \quad (5)$$

It is important to note that the perturbation is assigned a negative sign because of the decrease in SNR from the baseline to the Lombard condition. Additionally, during the baseline phase, there is no acoustic perturbation and, therefore, the SPL-pert value from Eq. (5) is zero. Similarly, the SNR-pert value during the recovery phase is also zero as the noise introduced in the Lombard phase is removed.

III. RESULTS

The SPL results for each condition and both participant groups are presented in Fig. 2. The mean SPL values for the baseline condition were 76.7 dB for the NPVH group and 79.9 dB for the typical voice group. Under the Lombard condition, these values increased to 81.9 dB for the NPVH group and 83.7 dB for the typical voice group. As a result of the increase in background noise from 35 to 80 dB, the SNR in the Lombard condition decreased by 45 dB for both groups, resulting in a SNR of −3.3 dB for the NPVH group and −0.1 dB for the typical voice group. This SNR reduction prompted compensatory behavior with SPL increases of +5.2 dB in the NPVH group and +3.8 dB in the typical voice group relative to their baseline conditions. Regarding the recovery conditions, the mean SPLs were 79.6 dB for the NPVH group and 79.9 dB for typical voice group. The participants with typical voices returned to baseline voice intensity levels, exhibiting a decrease of −5.2 dB. In contrast, participants with NPVH did not return to baseline voice intensity levels, showing a decrease of −2.3 dB when the noise in recovery condition is removed.

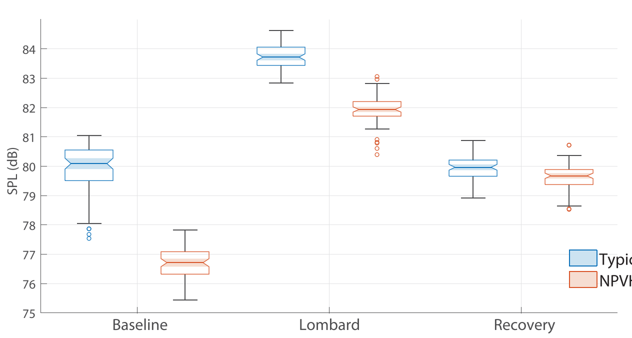

FIG. 2. (Color online) SPL boxplot for the three experimental conditions: (1) baseline (in quiet), (2) Lombard (in noise), and (3) recovery (in quiet after 5 min of rest) for participants with typical voices (blue) and participants with NPVH (orange).

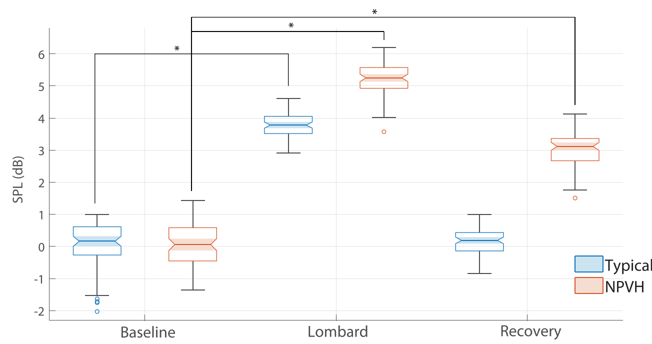

FIG. 3. (Color online) Boxplot of difference-SPL by participants with typical voices (blue) and participants with NPVH (orange) across the three experimental conditions: (1) baseline (in quiet), (2) Lombard (in noise), and (3) recovery (in quiet after 5 min of rest). Statistical difference is displayed with black lines (* $p < 0.05$).

Given that participants with NPVH began the experiment with a lower SPL than those with typical voice, a second analysis was conducted. In this analysis, we computed the difference between the Lombard condition and baseline condition as well as between the recovery and baseline conditions. For each participant, we calculated the mean SPL of the utterances corresponding to the baseline condition. Then, we computed the difference between this value and each SPL measurement for each repetition of the baseline, Lombard, and recovery conditions. This method is similar to that reported in a previous study (Castro *et al.*, 2022a) and used to enhance the comparison of SPL between both groups because of the difference in starting points. This process was performed individually for each participant. We refer to this measure as difference-SPL, and it is depicted in Fig. 3. The results of a two-way ANOVA are presented in Table I and Fig. 3.

The variation trial to trial was analyzed using the SimpleDIVA model to explore the dynamics among feedforward (FF), auditory gain (α_A), somatosensory gain (α_S), and learning rate (λ_{FF}). Given the baseline (or target) SPL, we calculated a SPL-pert of −0.5 and −0.67 for participants with typical voice and participants with NPVH, respectively. The model fit is depicted in Fig. 4, and the resulting parameters are shown in Table II.

IV. DISCUSSION

A. Speaking in noise and aftereffect

Our results showed that participants with typical voices and NPVH exhibited an increase in SPL of their voice

TABLE I. Statistical results for two-way ANOVA for difference-SPL.

Cases	Sum of squares[a]	Degrees of freedom	Mean square	F	p
Group	62.8	1	62.8	23.9	<0.001
Condition	412.6	2	206.3	78.4	<0.001
Group ∗ condition	10.6	2	20.3	7.72	<0.001
Residuals	300.0	114	2.63	—	—

[a]Type III sum of squares.

https://doi.org/10.1121/10.0034544

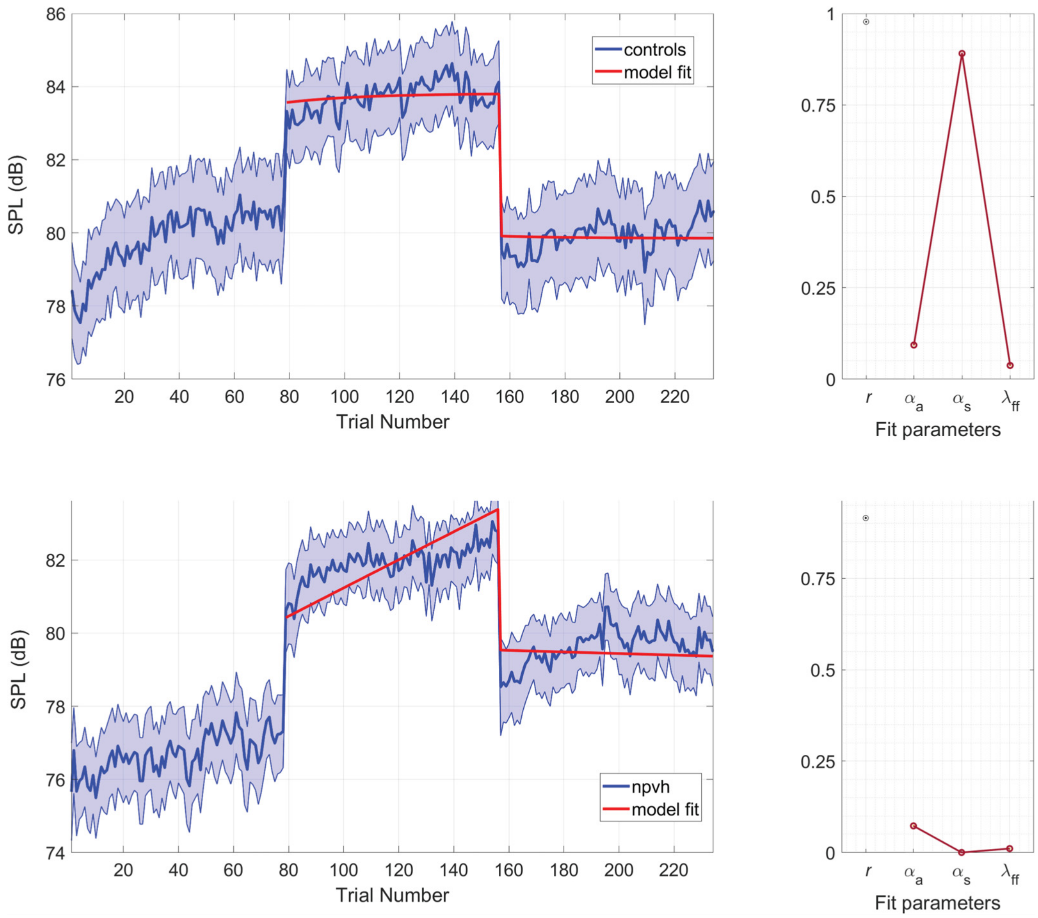

FIG. 4. (Color online) The results of model for participants with typical voices (top) and participants with NPVH (bottom) across the three experimental conditions: (1) baseline (in quiet), (2) Lombard (in noise), and (3) recovery (in quiet after 5 min of rest).

during the Lombard condition in response to the masking noise. This increase in SPL that results from masking noise is consistent with previous studies of the LE conducted via headphones (Castro *et al.*, 2022b; Garnier *et al.*, 2010; Meekings *et al.*, 2016; Stowe and Golob, 2013).

The results illustrate that the compensatory response to noise perturbation offers valuable insights into the relationship between SNR variation and vocal adaptation across conditions. As displayed in Fig. 2, participants with typical voices exhibited a compensatory increase of +3.8 dB when the SNR was −0.1 dB in the Lombard condition, resulting in a final SNR of 3.7 dB. In contrast, participants in the NPVH group showed a larger compensatory increase of +5.2 dB with an initial SNR of −3.3 dB in the Lombard condition, which led to a smaller SNR of 1.9 dB. The observed difference between the two groups can be attributed to the fact that NPVH participants began the experiment with lower vocal intensity during the baseline condition (in quiet) compared to the control group. As a result, NPVH participants need to exert greater effort to achieve a functional SNR in a noisy environment, aligning with previous findings that link NPVH to increased effort in voice production (Espinoza *et al.*, 2017). Despite this greater effort, their resulting SNR remains lower than that of the control group. This suggests that the NPVH group may opt for a slightly reduced but functional SNR to minimize the additional effort required.

TABLE II. Optimized parameters estimated by the SimpleDIVA model using acoustic SPL as experimental stimuli (ρ, Pearson's correlation; RMSE, root mean square error).

	SPL-pert (–)[a]	α_A	α_S	λ_{FF}	ρ	RMSE[b]
Controls	–0.50	0.09	0.89	0.04	0.98	0.07
NPVH	–0.67	0.07	0.00	0.01	0.91	0.06

[a]Downward perturbation to SPL-pert according to SPL targets.
[b]Fit to individual-subject/sample data.

The varying responses to the same noise levels highlight the importance of baseline vocal intensity in determining the LE. This is particularly evident during the recovery phase, where NPVH participants do not return to their original baseline intensity but instead maintain a louder voice. Our analysis showed that both groups decreased vocal intensity after 5 min of rest following noise removal. However, participants with NPVH did not return to baseline levels, showing a significant difference between the baseline and recovery conditions. Previous studies suggest that extended exposure to the LE in individuals with NPVH may create a mismatch between predicted vocal intensity (feedforward commands) and real-time auditory feedback because of the higher background noise. This discrepancy triggers corrective motor commands, increasing vocal intensity to align with predictions. Over time, speakers adapt to the noisy environment, updating their feedforward commands. When the noise is removed, the feedforward command prediction once again does not match their intensity needs, requiring time to reestablish appropriate control, leading to the observed aftereffect (Castro *et al.*, 2022b). It appears that 5 min of rest is sufficient for updating the feedforward system in individuals with typical voices. However, for those with NPVH, this period may be inadequate to fully reset their feedforward commands, resulting in persistently elevated SPL levels even after the noise is removed. SPL levels during recovery were similar between the two groups. This behavior is concordant with the observed in previous studies (Castro *et al.*, 2022b).

B. Using SimpleDiva model for exploring LE adaptation

We used the SimpleDIVA model to explore the adaptation to speaking in noise (LE) in participants with typical voices and those with NPVH. The results indicated that participants with NPVH have a lower feedforward learning rate and somatosensory gain compared to individuals with typical voices. Both groups exhibited similar values of auditory gain. Our results showed that participants with NPVH exhibited lower learning rate values than individuals with typical voices. This could be associated with a lower capacity of the participants with NPVH to adapt to variations in the acoustic background conditions, requiring more time to update the feedforward process compared to that for the control group.

According to the somatosensory gain estimated by the model, it is important to mention that previous studies using the SimpleDIVA model to explore sensorimotor adaptation with masking noise and formant shift propose that masking noise interferes with auditory feedback but does not eliminate somatosensory feedback. Thus, the somatosensory feedback controller attempts to move the vocal tract back toward its pre-perturbation configuration. The resulting corrective movements generated by the somatosensory feedback controller lead to updating of the feedforward commands, resulting in the de-adaptation evident in the experimental data and model fit (Ballard *et al.*, 2018; Kearney *et al.*, 2020). In our results, participants with typical voices showed an increase in somatosensory gain, suggesting that speakers use somatosensory feedback gain at the expense of auditory feedback gain when speaking in noise for updating the feedforward. This is consistent with previous studies with formants in noise, which presented higher levels of somatosensory gain compared to simulations using perturbation of formant without the use of masking noise during speech. Apparently, the use of masking noise as auditory input increases the somatosensory feedback gain estimated by the model (Kearney *et al.*, 2020).

Our analysis of the results using the SimpleDIVA model suggests that individuals with NPVH rely more heavily on auditory information than on somatosensory feedback for updating feedforward commands when speaking under masking noise Also, the participants with NPVH exhibit difficulties in updating the feedforward commands, characterized by lower adaptation under speaking in noise and an increase in the aftereffect phenomenon. This reinforces the idea that individuals with NPVH may require more time to adapt to variations in auditory environments than individuals with typical voices. Nevertheless, it is necessary to consider that the increase in the somatosensory feedback gain shown by the control group may be influenced by the optimization methods of the model as it was originally developed for auditory perturbations characterized by artificial manipulations such as formant and pitch shifts.

The findings of this study may have clinical implications. First, the analysis of SimpleDiva model indicates that participants with NPVH use less somatosensory feedback information during the vocal sensorimotor adaptation process compared to participants with typical voice. This suggests that individuals with NPVH primarily rely on auditory information to adjust their voice production in response to speaking in noisy environments. Moreover, the variation in the SNR observed from baseline condition to Lombard condition in participants with NPVH indicates the need for greater effort to achieve a functional SNR while speaking in noise. This behavior may be associated with an increase in vocal effort. This heightened effort could adversely affect voice production, potentially leading to elevated levels of vocal fatigue. Given that exposure to noisy environments is a common aspect of daily life, it is crucial to consider therapeutic strategies aimed at reducing a possible vocal fatigue caused by speaking in noise. This includes the development of new therapeutic approaches that stimulate the use of somatosensory feedback in patients with NPVH.

Another aspect with potential clinical relevance is the prolonged aftereffect observed in participants with NPVH. In this context, a 5-min rest period may be sufficient for participants with typical voices to return to baseline conditions. However, participants with NPVH may require more time to update their feedforward commands once the noise has been removed and the environment is quiet again. Furthermore, both groups of participants exhibited similar values for the voice intensity and SNR following the removal of noise in the recovery condition. This finding can be interpreted in two ways. First, the background noise may lead to persistently elevated vocal effort in participants with NPVH, which continues even after the noise has been removed, potentially contributing to vocal fatigue. Alternatively, the aftereffect of speaking in noise might actually benefit individuals with NPVH; in this scenario, speaking under masking noise could serve as a clinical strategy to disrupt VH and assist in reestablishing an adequate SNR in patients with NPVH. However, further studies are needed to explore the potential utility of the aftereffect phenomenon that results from speaking in noise as a therapeutic tool.

Finally, the speaking in noise or LE is an adaptive mechanism to maintain the oral communication shared by several animal species such as birds, mammals, or even fish. Also, it is modulated by linguistic and communicative contexts, demonstrating its evolutionary importance. Therefore, the LE should not be considered exclusively as an auditory feedback perturbation. Linguistic and communicative aspects should be explored in future studies with individuals with NPVH. Similarly, the value estimates from the model, learning rate, somatosensory feedback gain, and auditory feedback gain should be complemented by neurophysiological study tools, such as functional magnetic resonance imaging or electroencephalography, in future studies for increasing the knowledge in the motor control of voice under different auditory environments in individuals with NPVH.

V. CONCLUSION

The aim of this study was to investigate the role of auditory feedback, somatosensory feedback, and feedforward processes of the LE in individuals with typical voice and those with NPVH. The results of our experiments showed that speaking in noise produced a compensatory response in participants with typical voices and participants with NPVH by increasing their voice intensity. Furthermore, this compensation was found to be related to the variation in SNR caused by changes in the background noise. However, individuals with NPVH exhibited a higher aftereffect from speaking in noise, even 5 min after the noise was removed. In this context, the SimpleDIVA model predicted differences in feedforward and somatosensory gain for both groups. In contrast, participants with NPVH showed lower values for feedforward learning rate and somatosensory feedback gain, but they showed similar levels for auditory feedback gain. These results suggest that participants with NPVH rely more on auditory feedback information to update the feedforward process when speaking in noise compared to participants with typical voices.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. National Institutes of Health (NIH), National Institute on Deafness and Other Communication Disorders Grant No. P50 DC015446 (M.Z.); Agencia Nacional de Investigación y Desarrollo de Chile (ANID) Grant Nos. BASAL AFB240002 (A.W. and M.Z.), FONDECYT 1230828 (M.Z.), and ANILLO ACT210053 (A.W); Department of Speech and Language Pathology, Universidad de Valparaíso and Ph.D. Program in Health Science and Engineering Universidad de Valparaíso, Valparaíso, Chile. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

AUTHOR DECLARATIONS

Conflict of Interest

A.W. and M. Z. have a financial interest in Lanek SPA, a company focused on developing and commercializing biomedical devices and technologies. Their interests have been reviewed and are managed by Universidad Técnica Federico Santa María in accordance with its conflict-of-interest policies.

Ethics Approval

The consent form was approved by the Research and Ethics Committee of the Faculty of Medicine, Universidad de Valparaiso, Chile, in accordance with assessment statement code 52015, and it complies with national guidelines for research involving human subjects as well as the Declaration of Helsinki.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L. F., Noordzij, J. P., and Stepp, C. E. (**2021**). "Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders," Sci. Rep. **11**, 13123.

Alghamdi, N., Maddock, S., Marxer, R., Barker, J., and Brown, G. J. (**2018**). "A corpus of audio-visual Lombard speech with frontal and profile views," J. Acoust. Soc. Am. **143**(6), EL523–EL529.

Aronson, A. E., and Bless, D. M. (**2009**). *Clinical Voice Disorders*, 4th ed. (Thieme Medical, New York).

Ballard, K. J., Halaki, M., Sowman, P., Kha, A., Daliri, A., Robin, D. A., Tourville, J. A., and Guenther, F. H. (**2018**). "An investigation of compensation and adaptation to auditory perturbations in individuals with acquired apraxia of speech," Front. Hum. Neurosci. **12**, 510.

Behroozmand, R., and Sangtian, S. (**2018**). "Neural bases of sensorimotor adaptation in the vocal motor system," Exp. Brain Res. **236**, 1881–1895.

Bottalico, P. (**2018**). "Lombard effect, ambient noise, and willingness to spend time and money in a restaurant," J. Acoust. Soc. Am. **144**(3), EL209–EL214.

Castro, C., Prado, P., Espinoza, V. M., Testarts, A., Marfull, D., Manriquez, R., Stepp, C. E., Mehta, D. D., Hillman, R. E., and Zañartu, M. (**2022a**). "Lombard effect in individuals with nonphonotraumatic vocal hyperfunction: Impact on acoustic, aerodynamic, and vocal fold vibratory parameter," J. Speech, Lang., Hear. Res. **65**(8), 2881–2895.

Castro, C., Rivera, L. Z., Prado, P., Cuadros, J., Cortés, J. P., Weinstein, A., Espinoza, V., and Zañartu, M. (**2022b**). "Acoustic and neurophysiological aspects of Lombard effect," bioRxiv: https://www.biorxiv.org/content/early/2022/10/03/2022.09.30.510383.

Cuadros, J., Z-Rivera, L., Castro, C., Whitaker, G., Otero, M., Weinstein, A., Martínez-Montes, E., Prado, P., and Zañartu, M. (**2023**). "DIVA meets EEG: Model validation using formant-shift reflex," Appl. Sci. **13**(13), 7512.

Darius, S., Voigt-Zimmermann, S., and Böckelmann, I. (**2023**). "Effects of occupation-specific vocal stress on the mental health of day care teachers," Folia Phoniatr. Logop. **75**(5), 306–315.

Demmink-Geertman, L., and Dejonckere, P. H. (**2002**). "Nonorganic habitual dysphonia and autonomic dysfunction," J. Voice **16**(4), 549–559.

Espinoza, V. M., Zañartu, M., Van Stan, J. H., Mehta, D. D., and Hillman, R. E. (**2017**). "Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction," J. Speech., Lang., Hear. Res. **60**(8), 2159–2169.

Fujiki, R. B., and Sivasankar, M. P. (**2017**). "A review of vocal loading tasks in the voice literature," J. Voice **31**, 388.e33–388.e39.

Garnier, M., Henrich, N., and Dubois, D. (**2010**). "Influence of sound immersion and communicative interaction on the Lombard effect," J. Speech, Lang., Hear. Res. **53**(3), 588–608.

Guenther, F. H., Hampson, M., and Johnson, D. (**1998**). "A theoretical investigation of reference frames for the planning of speech movements," Psychol. Rev. **105**(4), 611–633.

Hillman, R. E., Stepp, C. E., Stan, J. H. V., Zañartu, M., and Mehta, D. D. (**2020**). "An updated theoretical framework for vocal hyperfunction," Am. J. Speech. Lang. Pathol. **29**, 2254–2260.

Hodgson, M., Steininger, G., and Razavi, Z. (**2007**). "Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments," J. Acoust. Soc. Am. **121**(4), 2023–2033.

Houde, J. F., and Jordan, M. I. (**1998**). "Sensorimotor adaptation in speech production," Science **279**(5354), 1213–1216.

Houde, J. F., and Nagarajan, S. S. (**2011**). "Speech production as state feedback control," Front. Hum. Neurosci. **5**, 82.

Huilgol, S., Baik, J., and Shattuck-Hufnagel, S. (**2019**). "A framework for labeling speech with acoustic cues to linguistic distinctive features," J. Acoust. Soc. Am. **146**(2), EL184–EL190.

Junqua, J. (**1993**). "The Lombard reflex and its role on human listeners and automatic speech recognizers," J. Acoust. Soc. Am. **93**(1), 510–524.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (**1977**). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J. Acoust. Soc. Am. **61**(5), 1337–1351.

Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S. E., Chao, S. C., Murray, E. S. H., Scott, T. L., and Guenther, F. H. (**2020**). "A simple 3-parameter model for examining adaptation in speech and voice production," Front. Psychol. **10**, 2995.

Kennedy, J., and Eberhart, R. (**1995**). "Particle swarm optimization," in *Proceedings of ICNN'95—International Conference on Neural Networks*, 27 November–1 December, Perth, WA, Australia (IEEE, New York), Vol. 4, pp. 1942–1948.

Lamotte, A.-S., Essadek, A., Shadili, G., Perez, J.-M., and Raft, J. (**2021**). "The impact of classroom chatter noise on comprehension: A systematic review," Percept. Motor Skills **128**(3), 1275–1291.

Lane, H., and Tranel, B. (**1971**). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**(4), 677–709.

Lombard, E. (**1911**). "Le signe de televation de la voix" ("The sign of the elevation of the voice"), Ann. Maladies Oreille Larynx, Nez, Pharynx **37**, 101–119.

Lu, Y., and Cooke, M. (**2008**). "Speech production modifications produced by competing talkers, babble, and stationary noise," J. Acoust. Soc. Am. **124**(5), 3261–3275.

McKenna, V. S., Hylkema, J. A., Tardif, M. C., and Stepp, C. E. (**2020**). "Voice onset time in individuals with hyperfunctional voice disorders: Evidence for disordered vocal motor control," J. Speech., Lang., Hear. Res. **63**, 405–420.

Mealings, K., Maggs, L., and Buchholz, J. M. (**2024**). "The effects of classroom acoustic conditions on teachers' health and well-being: A scoping review," J. Speech, Lang., Hear. Res. **67**(1), 346–367.

Meekings, S., Evans, S., Lavan, N., Boebinger, D., Krieger-Redwood, K., Cooke, M., and Scott, S. K. (**2016**). "Distinct neural systems recruited when speech production is modulated by different masking sounds," J. Acoust. Soc. Am. **140**(1), 8–19.

Nahid, M., and Hodgson, M. (**2011**). "Prediction of optimal conditions for verbal-communication quality in eating establishments," J. Acoust. Soc. Am. **129**(4), 2005–2014.

Núñez-Batalla, F., Morato-Galán, M., García-López, I., and Ávila-Menéndez, A. (**2015**). "Validation of the Spanish adaptation of the consensus auditory-perceptual evaluation of voice (CAPE-V)," Acta Otorinolaringol. (English Ed.) **66**(5), 249–257.

Pick, H. L., Jr., Siegel, G. M., Fox, P. W., Garber, S. R., and Kearney, J. K. (**1989**). "Inhibiting the Lombard effect," J. Acoust. Soc. Am. **85**(2), 894–900.

Stepp, C. E., Lester-Smith, R. A., Abur, D., Daliri, A., Noordzij, J. P., and Lupiani, A. A. (**2017**). "Evidence for auditory-motor impairment in individuals with hyperfunctional voice disorders," J. Speech., Lang., Hear. Res. **60**, 1545–1550.

Stevens, K. N. (**2002**). "Toward a model for lexical access based on acoustic landmarks and distinctive features," J. Acoust. Soc. Am. **111**(4), 1872–1891.

Stowe, L. M., and Golob, E. J. (**2013**). "Evidence that the Lombard effect is frequency-specific in humans," J. Acoust. Soc. Am. **134**(1), 640–647.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., and Stokes, M. A. (**1988**). "Effects of noise on speech production: Acoustic and perceptual analyses," J. Acoust. Soc. Am. **84**(3), 917–928.

Tartter, V. C., Gomes, H., and Litwin, E. (**1993**). "Some acoustic effects of listening to noise on speech production," J. Acoust. Soc. Am. **94**(4), 2437–2440.

Tourville, J. A., and Guenther, F. H. (**2011**). "The diva model: A neural theory of speech acquisition and production," Lang. Cognit. Process. **26**(7), 952–981.

Xia, T., Raneses, J., Schmiesing, B., Garcia, R., Walding, A., DeMajo, R., Schulz, A., and Batterman, S. A. (**2024**). "How teacher behaviors and perceptions, air change rates, and portable air purifiers affect indoor air quality in naturally ventilated schools," Front. Public Health **12**, 1427116.

Xue, C., Kang, J., Hedberg, C., Zhang, Y., and Jiang, J. J. (**2019**). "Dynamically monitoring vocal fatigue and recovery using aerodynamic, acoustic, and subjective self-rating measurements," J. Voice **33**, 809.e11–809.e18.

Younes, M. N., Heikal, A. Z., Kotb, A. S., and Zohny, H. N. (**2021**). "Assessment of train noise at platforms in underground metro stations," Civil Environ. Eng. **17**(1), 125–138.