

Nonconscious Control of Voice Intensity During Vocalization

Honorata HAFKE

Adam Mickiewicz University
Institute of Acoustics
Umultowska 85, 61-114 Poznań, Poland
e-mail: honorata_hafke@o2.pl

(received June 30, 2009; accepted November 20, 2009)

There are two separate visual systems in the human brain. Evidence from studies on both the humans and other primates has shown that there is a distinction between vision for perception and vision for action, which is reflected in the organization of the visual pathways in the cerebral cortex of primates. In recent years, researchers have attempted to find a similar dissociation between action and perception in human audition. The hypothesis tested in this paper is that the voice intensity is tracked and controlled by an auditory motor system. The results of this control are used for nonconsciously correct the vocal production. To observe the dissociation between perception and motor control, a subliminal experimental situation was created, in which values below the perceptual threshold (values which were not processed through normal channels or apparatus of perception) were used. The hypothesis was that a subliminal modification of an auditory voice feedback would cause an appropriate correction as a response, even if this change was not actually perceived. Assuming that the auditory system functions in the same way as the visual one and processes the information vital for motor reactions in real time, a reaction that would compensate such a modification should be expected.

Keywords: vocalization control, vowels, speech, auditory feedback.

1. Introduction

Sensory channels analyze the incoming stimuli and, through specific algorithms, decode the information encapsulated inside, thus defining their meaning.

In case of both vision and audition, recent studies are aimed at discovering the algorithms used in the final stages of those processes to identify the objects and sound sources. The researchers are therefore trying to describe the complex process of perception. It is still not clear, how the brain interprets and uses that information to guide the actions. In visual information processing, thanks to

studies by GOODALE and MILNER [5, 13] as well as their associates [2, 4, 9, 16], a widely accepted theory of vision was established, describing the working mechanism of “visual brain in action”. The authors introduce definitions of the stream for perception and stream for action.

In vision, the motor stream is used to transform information that describe the features of objects, such as: shape, size, spatial orientation, in order to properly mediate visual control of skilled actions directed at them (extending hands towards the objects, setting palm aperture, etc.)

In audition, the definition of tasks that require analogous motor stream is not obvious. Studies [6, 7, 11] have shown the existence of a separate motor mechanism independent of the perception stream in the process of vocalization control. The peripheral mechanisms of pitch and amplitude control are well understood.

However, given our detailed understanding of the peripheral mechanisms of control, there is a paucity of information related to neural control mechanisms of voice parameters.

Researchers [3, 8] have demonstrated that participants can respond to an unanticipated perturbation in voice auditory feedback with a compensatory response. Data from these studies suggest that voice auditory feedback can be used to monitor and stabilize voice amplitude around a desired loudness.

The latencies of the responses, with a mean value between 150 and 300 ms, coupled with the fact that the direction of the responses is generally opposite to the stimulus direction, suggests that these responses are automatic in nature.

These observations suggest that the process could be analogous to reactions observed in experiments, focusing on changes in the voice fundamental frequency [6].

It is assumed that both, the mechanisms controlling the voice fundamental frequency and the sound pressure level, may function in a similar way.

The purpose of the present study was to investigate whether singers compensate the small perturbations in voice intensity feedback during vocalization, that were not perceived during the process of vocalization.

Presence of motor reactions in such situation could be another proof of the existence of a separate motor stream dedicated to the control of vocalization.

2. Method

2.1. Subjects

Ten adults (4 females and 6 males) between 21 and 28 years of age participated in the experiment. All the listeners qualified as having normal hearing, which was defined as the audiometric threshold of 20 dB hearing level, or better, for a range from 250 to 8000 Hz [1]. They reported no neurological defects and had no speech or voice disorders. All of them had experience in singing (most of them members of academic choir). The participants were seated in a sound-treated room.

2.2. Apparatus and stimuli

The experimental schema is presented in Fig. 1. The participants' voices were recorded with a Microtech Geffel M300 microphone, with a fixed 6 cm mouth-to-mike distance.

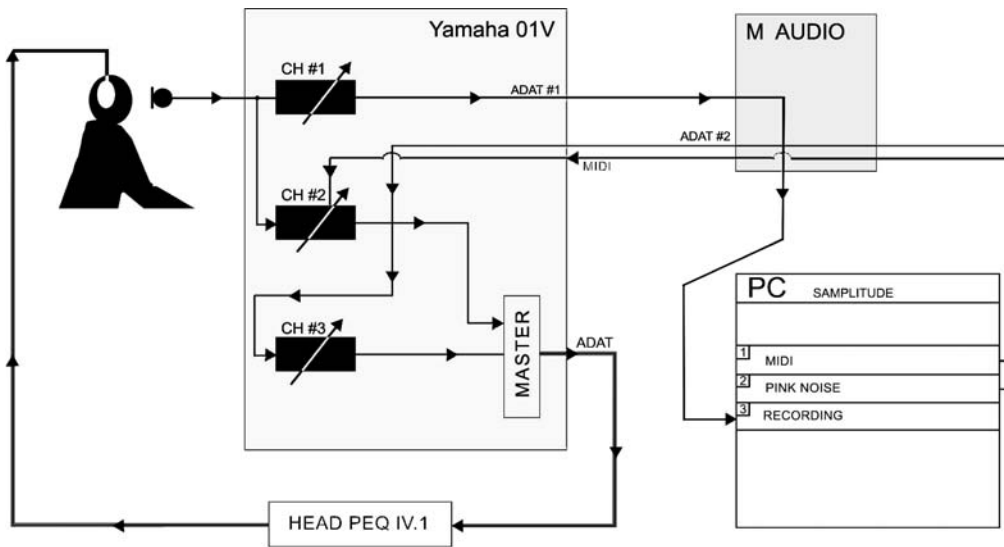


Fig. 1. Schema of the experimental setup.

The signal was then simultaneously directed through two ADAT outputs of a Yamaha 01V mixer to a recording device (PC class desktop computer) and (after introducing modifications in sound pressure level) to a PEQ IV.1 programmable equalizer. The signal was presented back to the subject via Sennheiser HD600 headphones, effectively creating the feedback loop. The regulation of audio amplification for feedback information was handled by automatic mechanism of the mixer, controlled by MIDI signal from Samplitude 8 programmable sequencer. An additional pink noise (60 dB SPL), generated by the recording device, was directed through the mixer back to the headphones (used in order to mask the bone-conducted auditory feedback).

2.3. Procedure

The experimental schema is presented in Fig. 2. Participants were instructed to vocalize the vowel /u/ for 5 seconds. The beginning and the end of a single vocalization was indicated by a short sound (a sine tone). Acoustic markers of different frequencies were used to mark the beginning and the end of vocalization.

Headphone amplification was set so that for the participant's preferred phonatory effort, the sound pressure level in headphones was 75 dB SPL.

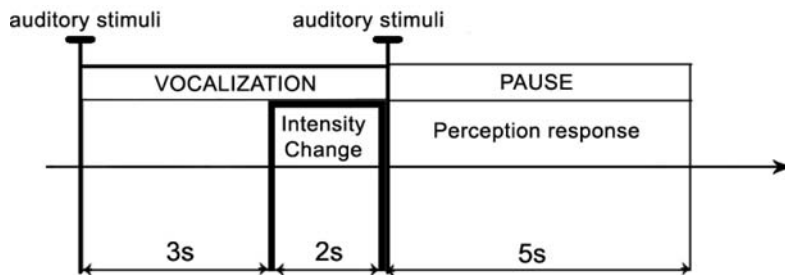


Fig. 2. Schematic illustration of the experimental procedure.

Listeners were asked to maintain constant voice level during the vocalization. Before each session, they were also presented with a tone of constant frequency (150 Hz for men, 240 Hz for women), and were asked to maintain the pitch of the vocalization at approximately the same level.

At the beginning of the third second of each vocalization, the voice sound pressure level in auditory feedback was increased or decreased by: 0.8, 1.6, 3.2, 5.2 dB respectively. The vocalization was recorded during the experiment, in order to determine reactions to perturbations introduced in auditory feedback.

Feedback modifications were presented randomly. Every change of sound pressure level was repeated 24 times. After each vocalization (Fig. 2) there was a pause, during which the participant decided whether or not she/he observed a change of intensity in her/his own voice.

2.4. Reaction check

All the data were analyzed using a dedicated application developed in MATLAB. A specific algorithm was used to determine motor reactions to changes in sound pressure level introduced in auditory feedback. For every vocalization, an average value of sound pressure level and standard deviation were calculated for a period of 300 ms before the onset of change in sound pressure level. Subsequently, a reaction check was performed. A positive reaction was defined as a deviation in sound pressure level, which had a latency of at least 50 ms, a magnitude of more than 2 SDs of the 300 ms pre-stimulus mean, and a duration of at least 50 ms in a maximum of 700 ms after the sound pressure level change. The application automatically determined the time when the averaged signal departed and reentered the 2 SD response criterion, and calculated the corresponding values in dB for valid responses only.

3. Results

As seen in Fig. 3, 82% of signals corresponded to the predefined criteria. The averaged results of the detection task and the averaged results of the motor re-

response are presented in Fig. 4a and Fig. 4b, respectively. Figure 4a represents the averaged results for the perception task of the experiment. Percentage values represent the number of detections of the sound pressure level changes. The averaged motor response values are presented in Fig. 4b. From the equation of the accumulative standard distribution curve, threshold values for the probability of 75% and 50% were calculated. The averaged threshold values for both the downward and upward changes were 1.9 dB (75%) and 1.3 dB (50%).

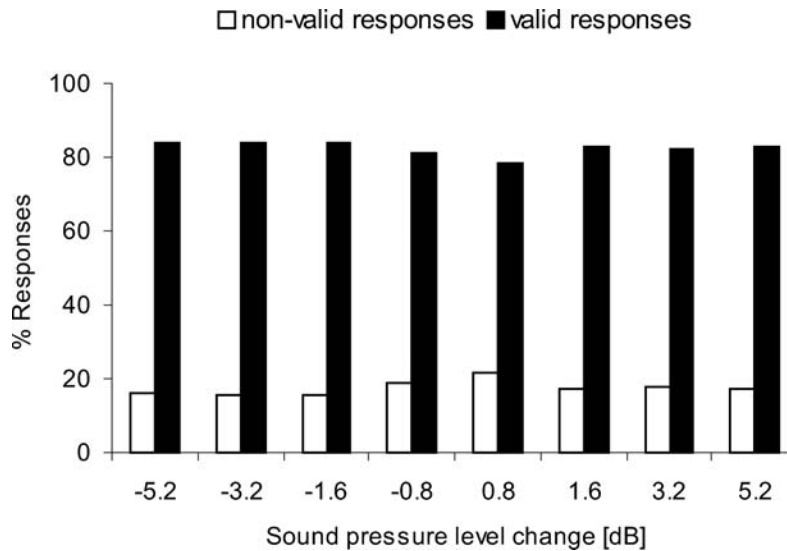


Fig. 3. Distribution of valid and non-valid responses for the predefined criteria. Valid response was defined as a deviation in sound pressure level, which had a latency of at least 50 ms, a magnitude of more than 2 SDs of the 300 ms pre-stimulus mean, and a duration of at least 50 ms in a maximum of 700 ms after the sound pressure level change.

In case of lowering the sound pressure level in auditory feedback, a motor reaction compensating the change (by increasing voice intensity) could be observed. An opposite reaction took place in case of increasing of the sound pressure level in auditory feedback. The proportionally largest compensation was noticed for the smallest changes, placed below the threshold in the perception part of the experiment.

The statistical analysis of the results was based on a two-way factorial ANOVAs with Tukey HSD posthoc tests. In case of the analysis of the main effects, value of the introduced change in auditory feedback was a statistically significant factor $F(7, 1789) = 79.18$; $p < 0.001$. Post-hoc (Tukey HSD) test for the level of significance $\alpha = 0.05$ showed, that there were statistically significant differences between reactions to upward (the increase in the sound pressure level) and downward (the decrease in the sound pressure level) changes in all cases (Table 1).

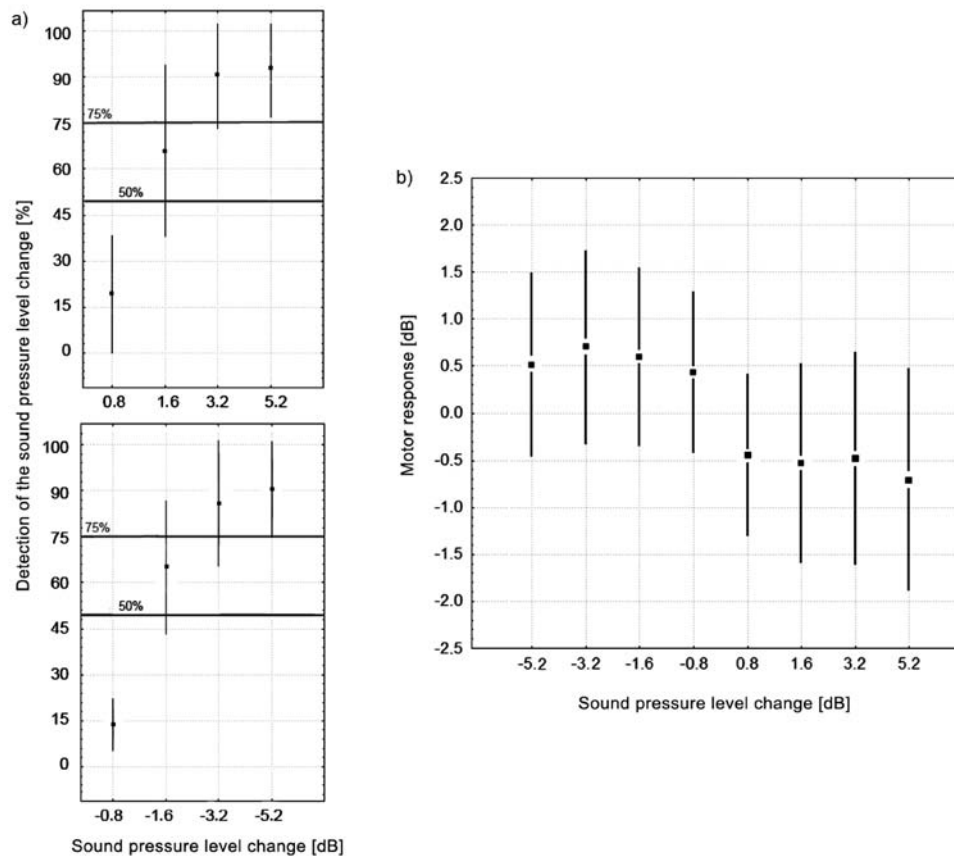


Fig. 4. The averaged results of the detection part of the experiment. Percentage values show the number of detections of the changes in sound pressure level (a). Horizontal lines mark the perception thresholds of 50% and 70%. Averaged motor reactions to the introduced changes in sound pressure level (b).

Table 1. Specification of results for post-hoc (Turkey HSD) comparisons. The first column shows the values of sound pressure level changes introduced in auditory feedback, for which the values of reactions were determined. The first row contains the numbers ascribed to those changes. Statistically significant differences are marked with bold font.

| Level Change | {1} | {2} | {3} | {4} | {5} | {6} | {7} | {8} |
|--------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| -5.2 {1} | | <i>0.49</i> | <i>0.99</i> | <i>0.99</i> | < 0.01 | < 0.01 | < 0.01 | < 0.01 |
| -3.2 {2} | <i>0.49</i> | | <i>0.95</i> | <i>0.07</i> | < 0.01 | < 0.01 | < 0.01 | < 0.01 |
| -1.6 {3} | <i>0.99</i> | <i>0.95</i> | | <i>0.64</i> | < 0.01 | < 0.01 | < 0.01 | < 0.01 |
| -0.8 {4} | <i>0.99</i> | <i>0.07</i> | <i>0.64</i> | | < 0.01 | < 0.01 | < 0.01 | < 0.01 |
| 0.8 {5} | < 0.01 | < 0.01 | < 0.01 | < 0.01 | | <i>0.98</i> | <i>1.00</i> | <i>0.12</i> |
| 1.6 {6} | < 0.01 | < 0.01 | < 0.01 | < 0.01 | <i>0.98</i> | | <i>1.00</i> | <i>0.62</i> |
| 3.2 {7} | < 0.01 | < 0.01 | < 0.01 | < 0.01 | <i>1.00</i> | <i>1.00</i> | | <i>0.27</i> |
| 5.2 {8} | < 0.01 | < 0.01 | < 0.01 | < 0.01 | <i>0.12</i> | <i>0.62</i> | <i>0.27</i> | |

4. Discussion

LOMBARD [10] was one of the first to investigate the topic of voice level control. Lombard effect is a phenomenon, in which speakers increase their vocal production in noisy environments. Studies [14] showed that this reaction cannot be suppressed by giving the speaker a specific set of instructions. More recent studies [3, 8] describe reactions to the perturbations of sound pressure level in voice auditory feedback. The results of these studies have demonstrated that subjects react to an unanticipated perturbation in sound pressure level feedback with an automatic compensatory response.

The authors [3] observed those reactions even for very small (1 dB) changes. The numerical values of reactions gathered in the present experiment were similar and supported these results.

However, none of the previous experiments tested whether the participants perceived changes in auditory feedback during the vocalization. It was assumed that motor reaction can only take place for consciously perceived changes. In this experiment, aside from determining the occurrence of reaction to changes in auditory feedback, the perception threshold for these changes was also calculated.

The following threshold values of perceptible level differences can be found in the literature [12]:

- 0.5–1 dB for wide-band noise (sound pressure level 20–100 dB above threshold),
- 1.5 dB for 20 dB SL; 0.7 dB for 40 dB SL; 0.5 dB for 80 dB SL for tones [15] in case of modulation detection.

ZWICKER [17, 18] stated that the listener is able to detect the change of sound pressure level, when the minimal change of stimulus is approximately 1 dB. In the conducted experiment it was determined that values of -0.8 dB and 0.8 dB were below the 50% threshold. Average percentage of detection for these values was below 20% (13% for -0.8 dB, 19% for 0.8 dB). Despite the fact that these changes could not be perceived, a reaction compensating the changes of sound pressure level in auditory feedback could be observed.

The results of the experiment support the hypothesis of a mechanism capable of nonconscious operation, that corrects small intensity perturbations present in the process of speech. Proportionally larger compensation was noticed for smaller changes than those for greater and more noticeable ones. This is consistent with the previously conducted researches [3]. A similar tendency can also be observed in studies focusing on the changes of voice fundamental frequency [6]. In case of consciously noticed changes, the same mechanism may be dependent on our will and intentions. A conscious decision could then be made to either compensate those changes or to ignore them. The obtained results suggest that (similar to control of the voice fundamental frequency) in case of motor control of voice intensity, a dissociation between perception and action for auditory information takes place.

Acknowledgment

This article is an extended version of the paper presented at the 56th Open Seminar on Acoustics – OSA2009, September 15–18 in Goniądz.

References

- [1] ANSI. ANSI S3.6-1996, *Specifications for Audiometers*, American National Standards Institute, New York 1996.
- [2] AGLIOTI S., DESOUZA J.F.X., GOODALE M.A., *Size-contrast illusions deceive the eye but not hand*, *Curr. Biol.*, **5**, 649–685 (1995).
- [3] BAUER J.J., MITTAL J., HAIN T.C., *Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude*, *J. Acoust. Soc. Am.*, **119**, 4, 2363–2371 (2006).
- [4] CULHAM J.C., DANCERT S.L., DESOUZA J.F.X., GATI J.S., MENON R.S., GOODALE M.A., *Visually-guided grasping produces fMRI activation in dorsal but not ventral stream brain areas*, *Exp. Brain Res.*, **153**, 180–189 (2003).
- [5] GOODALE M.A., HUMPHREY G.K., *The objects of action and perception*, *Cognition*, **67**, 181–207 (1998).
- [6] HAFKE H.Z., *Nonconscious control of fundamental voice frequency*, *J. Acoust. Soc. Am.*, **123**, 1, 273–278 (2008).
- [7] HAFKE H.Z., *Auditory information processing for sound perception and vocalization in case of Tritone Paradox*, *Archives of Acoustics*, **33**, 441–446 (2008).
- [8] HEINKS–MALDONADO T.H., HOUE J.F., *Compensatory responses to brief perturbations of speech amplitude*, *ARLO*, **6**, 3, 131–137 (2005).
- [9] KRÓLICZAK G., HEARD P., GOODALE M.A., GREGORY R.L., *Dissociation of perception and action unmasked by the hollow-face illusion*, *Brain Res.*, **1080**, 9–16 (2006).
- [10] LOMBARD E., *Le signe de le elevation de la voix*, *Larynx. Nez. Pharynx*, **37**, 101–119 (1911).
- [11] LOUI P., GUENTHER F.H., MATHYS C., SCHLAUG G., *Action perception mismatch in tone-deafness*, *Curr. Biol.*, **18**, 331–332 (2008).
- [12] MILLER G.A., *Sensitivity to changes in the intensity of white noise and its relation to masking and loudness*, *J. Acoust. Soc. Am.*, **61**, 89–108 (1947).
- [13] MILNER A.D., GOODALE M.A., *The Visual Brain in Action*, Oxford University Press, Oxford 1995.
- [14] PICK H.L., SIEGEL G.M., FOX P.W., GARBER S.R., KEARNEY J.K., *Inhibiting the Lombard effect*, *J. Acoust. Soc. Am.*, **85**, 2, 894–900 (1989).
- [15] RIESZ R.R., *Differential intensity sensitivity of the ear for pure tones*, *Physical Reviews*, **31**, 867–875 (1928).
- [16] WESTWOOD D.A., GOODALE M.A., *Perceptual illusion and the real-time control of action*, *Spatial Vision*, **16**, 243–254, (2003).
- [17] ZWICKER E., *Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs*, *Acustica*, **2**, 125–133 (1956).
- [18] ZWICKER E., *Masking and psychological excitation as consequences of the ear's frequency analysis*, [in:] *Frequency Analysis and Periodicity, Detection in Hearing*, PLOMP R., SMOORENBURG G.F. [Eds.], Sijthoff, Leiden 1970.