PAN
ARCHIVES OF ACOUSTICS
Vol. **38**, No. 3, pp. 375–381 (2013)

DOI: 10.2478/aoa-2013-0044

Pitch Processing of Speech: Comparison of Psychoacoustic and Electrophysiological Data

Małgorzata WRZOSEK[(1)], Justyna MACULEWICZ[(1)], Honorata HAFKE-DYS[(2)], Agnieszka NOWIK[(3)], Anna PREIS[(2)], Grzegorz KROLICZAK[(4)]

[(1)] *Institute of Psychology, Adam Mickiewicz University*
Szamarzewskiego 89, 60-568 Poznań, Poland; e-mail: wrzosek.gosia@gmail.com

[(2)] *Institute of Acoustics, Adam Mickiewicz University*
Umultowska 85, 61-614 Poznań, Poland

[(3)] *Department of Biophysics, Poznań University of Medical Sciences*
Fredry 10, 61-701 Poznań, Poland

[(4)] *Institute of Psychology, Cognitive Science Program, Head, Action, Cognition Laboratory, Adam Mickiewicz University*
Szamarzewskiego 89B, 60-568 Poznań, Poland

(*received September 30, 2012; accepted June 14, 2013*)

The present study consisted of two experiments. The goal of the first experiment was to establish the just noticeable differences for the fundamental frequency of the vowel /u/ by using the 2AFC method. We obtained the threshold value for 27 cents. This value is larger than the motor reaction values which had been observed in previous experiments (e.g. 9 or 19 cents). The second experiment was intended to provide neurophysiological confirmation of the detection of shifts in a frequency, using event-related potentials (ERPs). We concentrated on the mismatch negativity (MMN) – the component elicited by the change in the pattern of stimuli. Its occurrence is correlated with the discrimination threshold. In our study, MMN was observed for changes greater than 27 cents – shifts of ±50 and 100 cents (effect size – Cohen's $d = 2.259$). MMN did not appear for changes of ±10 and 20 cents. The results showed that the values for which motor responses can be observed are indeed lower than those for perceptual thresholds.

Keywords: pitch processing of speech, auditory perception-motor dissociation, mismatch negativity.

1. Introduction

The idea that there is a co-existence of two processing pathways in the visual cortex was first postulated in 1969 by Schneider. Thirteen years later, the hypothesis was developed by Ungerleider and Mishkin who consolidated the division of the visual processing systems into the ventral pathway – *what*, responsible for identifying visual objects, and the dorsal pathway – *where*, responsible for locating the stimulus (Ungerleider, Mishkin, 1982). The famous study of Patient D.F., a woman with damage in the brain region called the lateral occipital complex (LOC), enriched this attitude towards vision, in particular with regard to the visually guided action (Milner, Goodale, 1995; Goodale, Milner, 2004).

Growing interest in the functional organization of the auditory cortex has led to increased interest in whether or not there are two processing pathways in the auditory modality. Some researchers decided to try to find auditory systems similar to those found in the visual domain. Despite the consensus on what tasks the ventral pathway performs (involvement in the perception of sound source attributes), the role of the dorsal pathway is still controversial, dividing the researchers into two groups – those who view the pathway as being involved in the perceptual location of sounds and those who view the pathway as a functional auditory system responsible for motor activity. Researchers from the latter group claim that the dorsal pathway does more than merely localize sound sources and is in fact involved in the planning and output of motor response (in speech or body movements). According to the present state of research, this function is mainly carried out in speech (Hickok, Poeppel, 2004; Warren *et al.*, 2005; Hafke, 2008; 2009) or rhythm control (Repp, 2000; 2006).

Demonstrating the existence of the action pathway seems to be most effective in the case of automatic responses to perceptually unnoticed stimuli or to perceptually undetected changes in the sound stimuli reaching our auditory system. A variety of studies have demonstrated that the introduction of changes to real time auditory feedback (to self-generated sounds during vocalization) elicits compensative reactions with a latency of 100–150 ms in listeners, with speech sounds (Xu *et al.*, 2004; Chen *et al.*, 2007), consonants (Donath *et al.*, 2002; Natke, Kalveram, 2001; Natke *et al.*, 2003), and vowels (Burnett *et al.*, 1998; Larson *et al.*, 1996; Larson, 1998; Burnett, Larson, 2002; Sivasankar *et al.*, 2005; Liu, Larson, 2007; Hafke, 2008). It has been shown that compensative reactions to shifts in fundamental frequency remain the same when the listeners follow instructions not to respond to changes in feedback (Hain *et al.*, 2000). In the aforementioned studies, various interval values were used to introduce changes in the voice fundamental frequency. The easily noticed shift of 100 cents (one semitone) was most commonly used (Natke, Kalveram, 2001; Burnett, Larson, 2002; Donath *et al.*, 2002; Sivasankar *et al.*, 2005). A compensative reaction to much smaller shifts in auditory feedback was demonstrated in more recent studies (Liu, Larson, 2007; Hafke, 2008; Jones, Keough, 2009). The authors managed to observe a complete compensation for shifts as small as 9 cents (Hafke, 2008) and 10 cents (Liu, Larson, 2007). The main question is whether these minor frequency changes were consciously perceived, or whether they were due to their location being below the perception threshold and processed by a separate system. Hafke's study (Hafke, 2008) checked if the changes by ±9, 19, 50 and 100 cents introduced in feedback were perceived by listeners. The perceptual threshold of 26 cents was established, indicating that shifts of ±9 and 19 cents were not consciously perceived, yet while motor reactions were observed for them. To confirm Hafke's hypothesis that some compensative reactions are not followed by a conscious detection of change, the authors decided to also test brain responses to changes of 10, 20, 50 and 100 cents in both directions (±) by using the mismatch negativity (MMN) component. The aim of this part of the study was to test whether this EEG component, which occurs when a change in the pattern is perceived (Näätänen *et al.*, 1978; 2007; Näätänen, Alho, 1995; Picton *et al.*, 2000), would appear for changes in the fundamental frequency of the recorded vowel /u/ which, according to Hafke, are not consciously perceived. Positive verification – lack of an MMN response to changes of 10 and 20 cents – may suggest the existence of a separate pathway in the auditory system, unrelated to the perception motor pathway and designed to unconsciously control self-generated sounds.

The authors decided to perform a combined experiment (with two groups of participants – one for the psychophysical part, and one for the electrophysiological part) in order to: (1) determine the threshold of discrimination for the fundamental frequency (psychoacoustic approach); and (2) test whether the MMN component would appear for shifts in the fundamental frequency of vowel /u/ which, according to Hafke, are not perceived (electrophysiological approach).

2. Psychophysical experiment

2.1. Participants

Twenty-four listeners (9 men, 15 women), aged between 19 and 28, took part in the psychoacoustic part of this study. All the listeners were classified as having normal hearing, with normal hearing defined as the audiometric threshold of 20 dB (or better) hearing level for the range from 250 Hz to 8000 Hz (American National Standards Institute [ANSI], 1996). No neurological defects, speech disorders, or voice disorders were reported.

2.2. Procedure

In the first part of this study, an adaptive two-alternative forced choice (2AFC) procedure was used (Kingdom, Prins, 2010; Blauert, 2012; Blauert, Jekosch, 2012). Due to the previous use of the vowel /u/ in order to determine motor reactions (Hafke 2008; Liu, Larson, 2007), recordings of this vowel used were again as the stimulus. The recordings were prepared in an acoustically treated lecture booth using a G.R.A.S. 40AN omnidirectional microphone placed 40 cm from the mouth. A 44,100 Hz sampling frequency with 24 bits resolution was chosen to record the signals, using an RME DIGI96 PRO audio interface. The phase vocoder based on algorithm written in the MATLAB environment (Laroche, Dolson, 1999) served as the modifier of the pitch of the speech material. Two different recordings of female voices (vowel /u/) with the frequency of 247 Hz were used. The recordings only varied in timbre in order to ensure that the listeners did not adapt to one specific timbre. The duration of all stimuli was 2 s (the fragment of vocalization with the smallest deviation). The interval between the presented sounds was set to 0.5 s. All recordings were provided with 50 ms Hanning amplitude ramps at the beginning and the end of each stimulus. In each trial, sound examples were randomly chosen. All the stimuli had the same fundamental frequency and an equal RMS level. Before each trial, one of the two speakers was randomly chosen (for the testing and reference sound). The probability of the reference sound to be presented was equal: either as the first or second of the two sounds. The task required the listeners to decide which of the two sounds had the higher frequency. A single training session of 20 trials was used with feedback (information whether or not their answers were correct) in order to familiarize the listeners with the signals used. During the experiment, no feedback was provided. Starting with the difference of 100 cents between the sounds in a pair, the "2-yes 1-no" variant of the 2AFC procedure was used. The choice of this method resulted in decreasing the difference between sounds after two consecutive correct answers and increasing after each incorrect answer. A discrimination threshold was established, corresponding to a 70.07% probability of being correct. Multiplicative steps were used in the procedure, with two values of 1.5 and 1.25, respectively. At the beginning of the experiment, changes were introduced with the bigger step, and after the second reversal any further changes were made with the smaller step. Until the 12th reversal was reached, trials for each of the threshold values were continued. A single threshold value was calculated for each subject as an average of the values at the 8 final reversals. Three adaptive runs served as the basis for the final threshold value. All the signals were rendered on a PC connected with an equalizer (HEAD acoustics PEQ V) and presented over Sennheiser HD600 headphones. This part of the study was conducted in an acoustically treated listening booth.

3. Electrophysiological experiment

3.1. Participants

The EEG part of the study included seven participants (3 males, 4 females) with the mean age of 24.2 years (SD = 1.9). All the listeners were classified as having normal hearing, defined as the audiometric threshold of 20 dB (or better) hearing level for the range from 250 Hz to 8000 Hz (ANSI, 1996). Results from one person were excluded from further analyzes due to an error in the data file. Eventually, five right-handed participants and one left-handed participant were included in the analysis. No neurological deficits or serious head injuries in the past were reported. Only one person had had a musical education. This part of the study was conducted at the Action and Cognition Laboratory at the Institute of Psychology at Adam Mickiewicz University in Poznan, with the approval of the Ethics Committee of the Institute of Psychology at Adam Mickiewicz University in Poznan.

3.2. Procedure

During the experiment, participants were sitting in chairs with armrests at a distance of 80 cm from the fixation point which was located at the eye level. In the experiment, there were 960 repetitions of 9 types of recordings of the vocalized vowel /u/ (the same stimuli provided by Hafke (2008)). The recordings were presented in random order to the listeners. The whole presentation consisted of 720 repetitions of standard stimuli with a fundamental frequency of 247 Hz and 240 deviant stimuli with positive and negative shifts in frequencies of 10, 20, 50 or 100 cents.

The stimulus duration was set to 250 ms, with the interval between stimuli of 140 ms. The experiment was controlled by the Presentation 0.52 (Neurobehavioral System) software and the sounds were presented binaurally through ER-2 Etymotic Research headphones. Participants were instructed to focus their attention on a fixation point.

Electrophysiological brain responses were collected using the BioSemi ActiveTwo system. The 64 electrodes with DC amplifiers were set in accordance with the 10–20 system (Jasper, 1958). The EEG signal was recorded with a sampling rate of 1024 Hz. The removal of artifacts related to eye-movement signals was based on electrooculography (EOG). The EOG signal was recorded bipolarly with electrodes placed at the outer left and right lateral canthus (hEOG), above and below (vEOG) the left eye. During acquisition, reference was averaged across all the channels.

The EEG analysis was performed using Brain Vision Analyzer 2 (Brain Products GmbH, Munich, Germany). At the beginning, the signal was re-referenced off-line into two electrodes placed at the left and right mastoid. The filter band-pass set to 0.1–30 Hz (12 dB octave/slope) was used. The registered EEG signal was divided into segments with a length of 400 ms (from 50 ms before stimulus onset to 100 ms after stimulus offset, with a stimulus which lasted 250 ms). Baseline correction was performed by averaging the signal within the interval of 50 ms before presentation of the stimulus. The time window for MMN was set for the 160–220 ms after stimulus onset (Luck, 2005).

4. Results

The results of the experiments are presented in Fig. 1. In the psychophysical part of the study, the average frequency threshold value amounts to 27 cents (Fig. 1a). For comparison, Fig. 1a additionally includes the values obtained in the previous experiment (Hafke, 2008).

The distribution of results for the female and male listener groups can be treated as normally distributed, as shown in the Shapiro-Wilk test. Statistical differ-

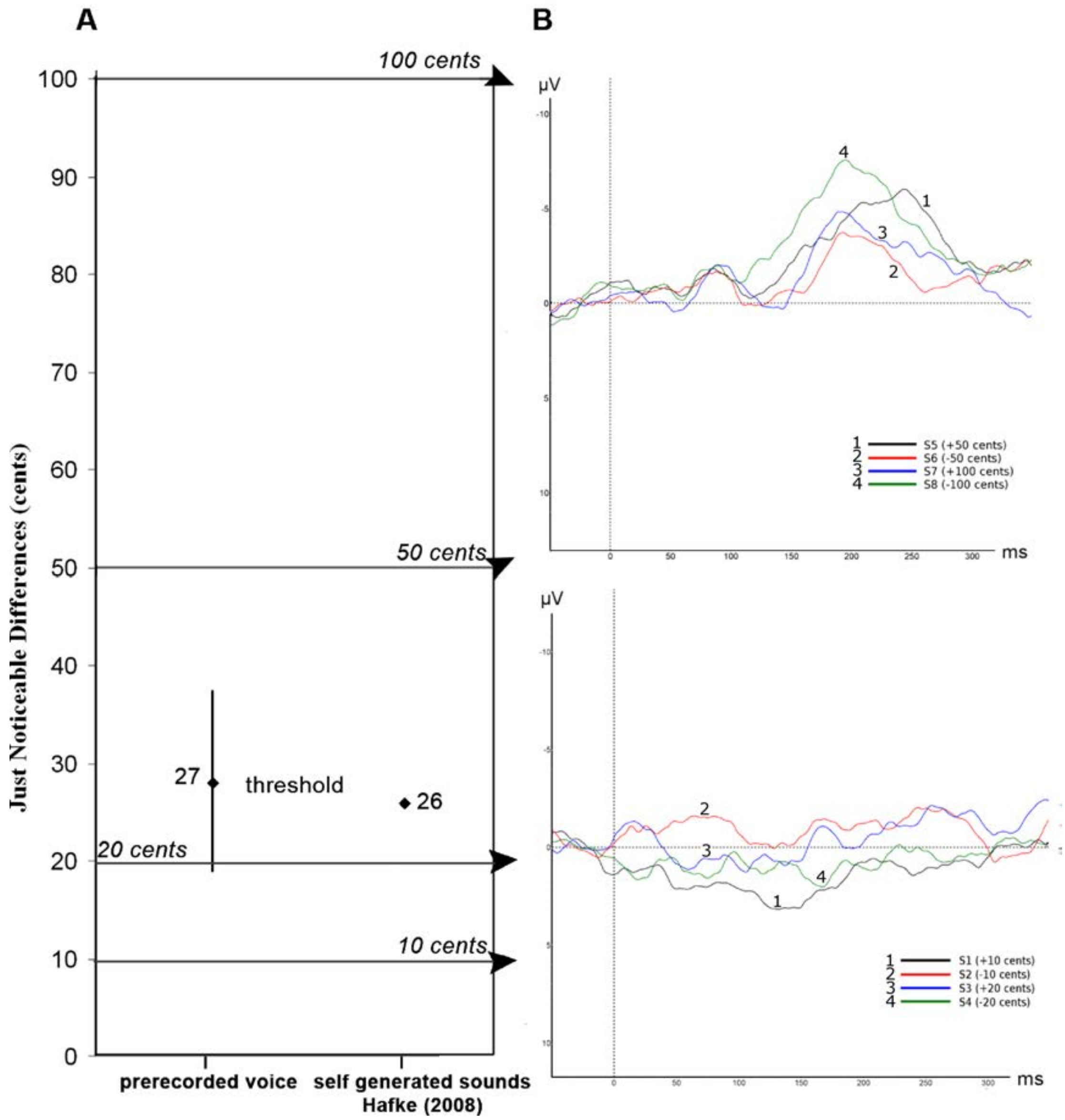

Fig. 1. Results: (**A**) The psychophysical part – the discrimination threshold established for 27 cents with confidence interval marked (value of 26 cents obtained in the previous experiment (Hafke, 2008) is on the right-hand side of the diagram); horizontal lines refer to shifts (of 10, 20, 50 and 100 cents) used in the EEG part – shifts of 50 and 100 cents and each shift larger than 27 cents are noticeable and should elicit MMN response; shifts of 10 and 20 cents are unnoticeable. (**B**) EEG part – deviant-minus-standard grand average difference waveforms at electrode Fz for noticeable (upper panel: ±50, 100 cents) and not noticeable shifts (lower panel: ±10 and 20 cents).

ences between the responses of the two gender groups ($F(1,\ 44) = 0.01$; $p = 0.93$) were not observed.

The occurrence of MMN for shifts of ±50 and ±100 cents is interpreted as a proof that changes in the frequency of vowel /u/ were detected. Any significant appearance of mismatch negativity was not recorded for undetected shifts (±10 and 20 cents) (Fig. 1b).

The obtained results were subjected to statistical analysis in IBM SPSS Statistics 20.0. The average MMN amplitude for stimuli containing a perceptually detectable shift was – 6.166 μV (SD = 3.172, SE = 0.647). The average amplitude for wave with the MMN-like latency for stimuli containing perceptually undetectable shifts (MMN was not elicited for these shifts) was – 1.582 μV (SD = 2.347, SE = 0.479). The average amplitude for the standard was – 0.732 μV (SD = 1.229, SE = 0.502). The average MMN or MMN-like latency for all stimuli was 195.245 ms (SD = 21.512, SE = 2.928).

Statistically significant differences ($F(8,\ 40) = 8.362$, $p < 0.001$) between the amplitudes of particular stimuli were observed, as shown in general linear model, repeated measures. Analysis of contrasts indicated that amplitudes for deviants with a change of +50, +100 and −100 cents comparing to standard stimuli were statistically significant (respectively:

$F(1, 5) = 24.475$, $p = 0.004$; $F(1, 5) = 15.387$, $p = 0.011$; $F(1, 5) = 27.373$, $p = 0.003$). When a 50 cent negative change was obtained, the result did not differ significantly from the standard ($F(1, 5) = 5.267$, $p = 0.07$). The same analysis of contrasts showed that changes of +10, −10, +20 and −20 cents did not elicit a statistically significant change in MMN amplitude when compared to the standard ($F(1, 5) = 0.050$, $p = 0.833$; $F(1, 5) = 2.415$, $p = 0.181$; $F(1, 5) = 5.448$, $p = 0.067$; $F(1, 5) = 0.742$, $p = 0.429$, respectively).

There were no significant differences in the amplitude of the mismatch negativity between the participants ($F(5, 48) = 0.936$, $p = 0.466$). The significance of differences in amplitudes of stimuli belonging to one from the three post-hoc one-way ANOVA groups – one containing shifts of ±10 and 20 cents (which did not elicit MMN response), the second containing shifts of ±50 and 100 cents (which elicited the MMN response) and the third containing standard stimuli (no shift) – was investigated. This relation was significant ($F(2, 51) = 21.250$, $p < 0.001$). A post-hoc two-sided Dunnett's Test was used to show that (1) these groups were significantly different from each other; (2) stimuli with detectable shifts (±50 and ±100 cents) differed significantly from the standard ($p < 0.001$); (3) deviants with shifts of 10 and 20 cents (undetectable shifts) did not differ significantly from the standard ($p = 0.646$); and (4) stimuli with detectable shifts differed significantly from stimuli with undetectable shifts ($p < 0.001$).

Cohen's d was calculated to demonstrate the effect size. In the case of the group containing undetectable shifts in comparison to the group containing standard stimuli, $d = 0.454$, while in the case of the group containing detectable shifts in comparison to the group containing standard stimuli, $d = 2.259$.

5. Discussion

In the previous experiment (2008), Hafke established for the fundamental frequency of self-generated sounds a perceptual threshold of 26 cents, using the method of the constant stimuli. In the psychoacoustic part of the present study there was no need for the subjects to produce speech sounds; participants were asked only to listen to speech that was generated externally. Although different experimental conditions and different psychophysical methods were used, the obtained perceptual threshold was almost the same.

The psychoacoustic part of the present study suggests that the motor reactions observed in the previous experiments were not associated with conscious perception of the introduced pitch changes. In summary, regardless of whether we look at the perception threshold for self-generated sounds (Hafke, 2008) or externally generated sounds (the present study), and regardless of the psychophysical method used, the perception threshold is always greater than the value for which motor reactions can be observed.

The independent confirmation was delivered by the second part of the present study. It turned out that the mismatch negativity (MMN), a component whose appearance is related to the detection of a perceptual deviation from a standard, was observed only for changes in the frequency of the vowel /u/ placed above the detection threshold established in the psychophysical part of the experiment. This result could be considered as a confirmation of the conclusion from the previous study (Hafke, 2008) about the occurrence of a compensatory motor response to non-perceived changes in the fundamental frequency. The mismatch negativity component was elicited by changes of 50 and 100 cents in both directions. For smaller shifts (10 and 20 cents), MMN was not observed.

In a study similar to the present one, Hawco *et al.* (2009) observed MMN in reaction to positive changes of 50, 100 and 200 cents in the vocalized vowel /a/. For the change of 25 cents, no MMN was observed, while a compensation response was noted. The present study constitutes an extension of Hawco's experiment and delivers a confirmation of the common parts. In the light of this study and the previous experiment (Hafke, 2008), the value of 25 cents can be considered subliminal to perception and classified as under the detection threshold. Table 1 presents the shifts used in the present study and by Hawco *et al.* (2009) with information whether the MMN component was present or not for these shift values. A compensative motor reaction was observed for all these shifts.

Table 1. The frequency shifts (in cents) used in the present study and by Hawco *et al.* (2009) – column 1, and the MMN responses for the fundamental frequency of vowel /u/ – column 2, and vowel /a/ – column 3, respectively.

Shifts (cents)	MMN – vowel /u/	MMN – vowel /a/
10	no MMN	–
20	no MMN	–
25	–	no MMN
50	MMN present	MMN present
100	MMN present	MMN present
200	–	MMN present

Note: '–' indicates changes not tested in particular experiment.

The results from the psychophysical part, as well as those of the electrophysiological part, suggest that the motor reactions observed in the previous studies (Liu, Larson, 2007; Hafke, 2008; Jones, Keough,

2009) occur for stimuli below the perceptual threshold. This may lead to the conclusion that there are two processing pathways co-existing in the auditory cortex.

The reader should take into account the fact that in the electrophysiological part of the study, the authors tested only 6 participants. This is not a common practice in EEG research. The authors consider this part of the present experiment as a preliminary study giving the insight into the problem and intend to extend this area in future research.

References

1. American National Standards Inst. (1996), ANSI S3.6-1996, *Specifications for Audiometers*, American National Standards Institute, New York.

2. Blauert J. (2012), *A perceptionist's view on psychoacoustics*, Arch. Acoust., **37**, 3, 365–372.

3. Blauert J., Jekosch U. (2012), *A Layer Model of Sound Quality*, J. Audio Eng. Soc., **60**, 1/2, 412.

4. Burnett T.A., Freedland M.B., Larson C.R., Hain T.C. (1998), *Voice F0 responses to manipulations in pitch feedback*, J. Acoust. Soc. Am., **103**, 6, 3153–3161.

5. Burnett T.A., Larson C.R. (2002), *Early pitch shift response is active in both steady and dynamic voice pitch control*, J. Acoust. Soc. Am., **112**, 1058–1063.

6. Chen S.H., Liu H., Xu Y., Larson C.R. (2007), *Voice F0 responses to pitch-shifted voice feedback during English speech*, J. Acoust. Soc. Am., **121**, 1157–1163.

7. Donath T.M., Natke U., Kalveram K.T. (2002), *Effects of frequency-shifted auditory feedback on voice F0 contours in syllables*, J. Acoust. Soc. Am., **111**, 357–366.

8. Goodale M.A., Milner M.A. (2004), *Sight Unseen: An Exploration of Conscious and Nonconscious Vision*, Oxford University Press, Oxford.

9. Hafke H.Z. (2008), *Nonconscious control of fundamental voice frequency*, J. Acoust. Soc. Am., **123**, 1, 273–278.

10. Hafke H.Z. (2009), *Nonconscious control of voice intensity during vocalization*, Archives of Acoustics, **34**, 4, 407–414.

11. Hain T.C., Burnett T.A., Kiran S., Larson C.R., Singh S., Kenney M.K. (2000), *Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex*, Exp. Brain Res., **130**, 133–141.

12. Hawco C.S., Jones J.A., Ferretti T.R., Keough D. (2009), *ERP correlates of online monitoring of auditory feedback during vocalization*, Psychophysiology, **46**, 1216–1225.

13. Hickok G., Poeppel D. (2004), *Dorsal and ventral pathways: a framework for understanding aspects of the functional anatomy of language*, Cognition, **92**, 1–2, 67–99.

14. Jasper H.H. (1958), *Report of the Committee on Methods of Clinical Examination in Electroencephalography*, Electroenceph. Clin. Neurophysiol., **10**, 370–371.

15. Jones J.A., Keough D. (2009), *The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing*, J. Acoust. Soc. Am., **126**, 2, 837–846.

16. Kingdom F.A.A., Prins N. (2010) *Psychophysics: A practical introduction*, Amsterdam: Academic Press.

17. Laroche J., Dolson M. (1999), *New Phase-Vocoder Techniques are Real-Time Pitch Shifting, Chorusing, Harmonizing, and Other Exotic Audio Modifications*, JAES, **47**, 11, 928–936.

18. Larson C.R. (1998), *Cross-modality influences in speech motor control: The use of pitchshifting for the study of F0 control*, J. Commun. Disorders, **31**, 489–503.

19. Larson C.R., White J.P., Freedland M.B., Burnett T.A. (1996), *Interactions Between Voluntary Modulations and Pitch-Shifted Feedback Signals: Implications for Neural Control of Voice Pitch*, [in:] *Vocal Fold Physiology: Controlling Complexity and Chaos*, P.J. Davis and N.H. Fletcher [Eds.], pp. 279–289, Singular, San Diego.

20. Liu H., Larson C.R. (2007), *Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex*, J. Acoust. Soc. Am., **122**, 6, 3671–3677.

21. Luck S. (2005), *An Introduction to the Event-Related Potential Technique*, MIT Press, Massachusetts Institute of Technology.

22. Milner A.D., Goodale M.A. (1995), *The Visual Brain in Action*, Oxford University Press, Oxford.

23. Natke U., Donath T.M., Kalveram K.T. (2003), *Control of voice fundamental frequency in speaking versus singing*, J. Acoust. Soc. Am., **113**, 3, 1587–1593.

24. Natke U., Kalveram K.T. (2001), *Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables*, J. Speech Lang. Hear. Res., **44**, 577–584.

25. Näätänen R., Gaillard A.W.K., Mäntysalo S. (1978), *Early selective-attention effect on evoked potential reinterpreted*, Acta Psychologica, **42**, 313–329.

26. Näätänen R., Alho K. (1995), *Mismatch negativity the unique measure of sensory processing in audition*, International Journal of Neuroscience, **80**, 317–337.

27. Näätänen R., Päävilainen P., Rinne T., Alho K. (2007), *The mismatch negativity (MMN) in basic re-*

search of central auditory processing: A review, Clin. Neurophysiol., **118**, 2544–2590.

28. Picton T.W., Alain C., Otten L., Ritter W., Achim A. (2000), *Mismatch negativity: Different water in the same river*, Audiology and Neuro-otology, **5**, 111–139.

29. Repp B.H. (2000), *Compensation for subliminal timing perturbations in perceptual-motor synchronization*, Psychological Research, **63**, 106–128.

30. Repp B.H. (2006), *Does an auditory perceptual illusion affect on-line auditory action control? The case of (de)accentuation and synchronization*, Exp. Brain Res., **168**, 493–504.

31. Sivasankar M., Bauer J.J., Babu T., Larson C.R. (2005), *Voice responses to changes in pitch of voice or tone auditory feedback*, J. Acoust. Soc. Am., **117**, 2, 850–857.

32. Ungerleider L.G., Mishkin M. (1982), *Two cortical visual systems. Analysis of visual behavior*, [in:] *Analysis of Visual Behavior*, Ingle D.J., Goodale M.A., Mansfield R.J.W. [Eds.], pp. 549–586, MIT Press, Cambridge.

33. Warren J.E., Wise R.J.S., Warren J.D. (2005), *Sounds do-able: auditory-motor transformations and the posterior temporal plane*, Trends in Neurosciences, **28**, 12, 636–643.

34. Xu Y., Larson C.R., Bauer J.J., Hain T.C. (2004), *Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences*, J. Acoust. Soc. Am., **116**, 2, 1168–1178.