

Assessing sensorimotor integration in adults who stutter by a behavioral task using perceptual adaptation of frequency-altered auditory feedback

Daichi Iimura^{1,2,*}, Nobuhiko Asakura³, Takafumi Sasaoka⁴ and Toshio Inui⁵

¹Faculty of Rehabilitation, Kawasaki University of Medical Welfare,
288 Matsushima, Kurashiki, 701-0193 Japan

²Graduate School of Comprehensive Human Sciences, University of Tsukuba,
1-1-1, Tennodai, Tsukuba, 305-8577 Japan

³Center for Mathematical Modeling and Data Science, Osaka University,
1-3 Machikaneyama-cho, Toyonaka, 560-8531 Japan

⁴Brain, Mind, and KANSEI Sciences Research Center, Hiroshima University,
1-2-3, Kasumi, Minami-ku, Hiroshima, 734-8553 Japan

⁵Department of Psychology, Otemon Gakuin University,
2-1-15 Nishiai, Ibaraki, 567-8502 Japan

(Received 18 January 2020, Accepted for publication 21 February 2020)

Keywords: Stuttering, Sensorimotor integration, Auditory feedback, Frequency-altered feedback, Adaptation
PACS number: 43.66.-x [doi:10.1250/ast.41.780]

1. Introduction

Developmental stuttering is a fluency disorder characterized by frequent word or part-word repetition, prolongation, and silent blocks that disrupt the rhythmic flow of speech [1]. Classically, the characteristic breakdowns in speech fluency can be attributed to sensorimotor speech deficits in adults who stutter (AWS) [2–6]. This notion is supported by neuroimaging [7], behavioral [4–6], and modeling [2] approaches.

One behavioral method to investigate sensorimotor integration is the adaptation paradigm. Adaptation is the process of adjusting one's perception or actions into new situations [8], such as visuo-motor, auditory-motor [6], or auditory-perception [4]. In a previous study using simultaneous judgment of adaptation of delayed auditory feedback (DAF), we found that perception in AWS was more strongly influenced by DAF compared to that in adults who do not stutter (ANS) [4]. We also investigated the relationship between the degree of perceptual influence by DAF and the degree of consistency of simultaneous judgment that was presumed to be a source of overreliance on auditory feedback. Results showed that the more imprecise the simultaneous judgment, the more AWS relied on auditory feedback. These findings suggested overreliance on the timing process of auditory feedback information in AWS, and that such overreliance was due, in part, to reduced internal reliability (i.e., inconsistent simultaneous judgment) [4].

However, it remains unclear whether this atypical sensorimotor integration is specific to timing processes (i.e., DAF). As Tsakiris *et al.* [9] proposed that the internal prediction signals that anticipate the sensory consequences of a motor command could be separated by timing and other information, timing and other properties of the voice could be processed independently. While experiments using fundamental frequency-altered feedback (FAF) were focused on

AWS [10], to date, no study has been conducted using the perceptual adaptation paradigm. The aim of the present study was to elucidate whether the previous findings of DAF adaptation [4] are applicable to processing of adaptation of fundamental frequency (Fo). We used a behavioral experiment using FAF of Fo adaptation in both AWS and ANS.

2. Methods

2.1. Ethics statement

The ethics committee to which the first author belonged approved the experimental procedures in advance.

2.2. Participants

The AWS group consisted of 20 adults (two women) aged 19–33 years (mean age and standard deviation, 22.9 ± 3.8 years); the ANS group consisted of 20 adults (two women) aged 20–25 years (22.1 ± 1.5 years). No participant had a history of neurological problems or speech or language problems, except for stuttering in the AWS group.

2.3. Experimental settings

The experiment was conducted in a closed, soundproof room individually. The participants were seated in a chair and look at a display monitor. The experimental setting and design were similar to those in a previous study [4]. The display monitor (LL-T174-B, SHARP, Osaka, Japan) was placed on a desk in front of the participant. Participants were required to look at the display wearing headphones (HP-RX500, JVC, Kanagawa, Japan). Their voices were recorded by a microphone (ECM-G5M, SONY, Tokyo, Japan) and fed back on the headphones through an auditory effector (MX300, LEXICON, Waltham, MA) and audio mixer in real-time (802VLZ4, MACKIE, Woodinville, WA). The auditory effector was connected to the microphone to produce frequency changes (FCs) of Fo between their produced voice and their voice perception via auditory feedback. FCs were manipulated using MATLAB (2012a, MathWorks, Natick, MA) on a Windows PC (PRECISION T1600, DELL,

*e-mail: iimura@mw.kawasaki-m.ac.jp

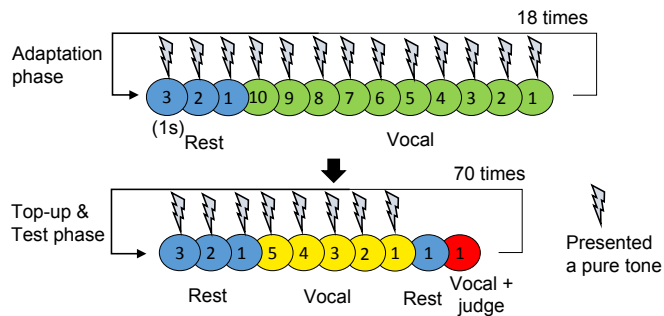


Fig. 1 Sequence of visual cues presented on the display in one block.

Tokyo, Japan). A MIDI controller connected between the PC and the effector, and this sent a command signal to the effector.

Pink noise ($L_{SPL} = 90$ dB) was constantly presented through the headphones during each block to mask the participants' voice via air duct sound and disrupt any potential additional auditory cues. The auditory feedback of the participants' voice and masking noise were composed through the audio mixer before being presented through the headphones. To induce perceptual adaptation, we did not set a particular sound level for their voice; instead, we confirmed with each participant that they could distinguish their own voice from the masking noise.

2.4. Procedures

We show the sequence of one block in Fig. 1. A visual cue with a blue, green, yellow, or red circle was successively presented at the center of the display for 1 s. Numbers on the colored circles were presented as a form of countdown (e.g., 3, 2, 1).

One block contained an "Adaptation phase" of inducing adaptation and a "Top-up & Test phase" of maintaining adaptation and judging. In the "Adaptation phase," three blue cues and 10 green cues were successively presented. This cycle was repeated 18 times. Blue cues denoted rest. When the green cue was presented, participants were required to say "ah." Since 10 green cues were presented, the participants said "ah" 10 times at intervals of 1 s per cycle. A pure tone (125 Hz for male participants and 240 Hz for female participants) was presented for 500 ms for each blue, yellow, and green cue through the headphone, and participants were asked to match their voice pitch to this model tone. Participants' voices were presented via the headphones with a constant FC during one block (the adapted FC: -4 , ± 0 , or $+4 \times 100$ cents compared to their F_0), inducing the frequency adaptation. In the "Top-up & Test phase," three blue cues, five yellow cues, a blue cue, and a red cue were successively presented, and this cycle was repeated 70 times. The yellow cues played the role of the green cues (the adapted FC was also the same). With the red cue, the participants were instructed to say "ah" and judge the difference between their voice expectation and actual auditory feedback at each FC (test FC of -3 , -2 , -1 , ± 0 , $+1$, $+2$ or $+3 \times 100$ cents). The test FCs in the 35 trials were in random order; every FC occurred five times in each block. The judgment was performed by pressing a keyboard

key (left arrow: high; up arrow: same; right arrow: low) located in front of the participants' right hand.

Every participant completed six blocks, with two blocks in each condition (adapted FC: -4 , ± 0 , or $+4 \times 100$ cents). Thus, there were 140 judgment trials in one condition and each of the seven FCs was repeated 20 times. The order of six blocks was counterbalanced across participants. There was a 5 min break for rest between blocks.

2.5. Analysis

The individual proportions of "high," "same" and "low" responses in each test FC (-3 , -2 , -1 , ± 0 , $+1$, $+2$ or $+3 \times 100$ cents) were calculated as a function of the FCs for each condition (adapted FC: -4 , ± 0 , or $+4 \times 100$ cents). In total, each participant completed six blocks (three conditions), and as a result, there were 20 judgments in each test FC in each adapted FC condition. A cumulative Gaussian psychometric function was then fitted by Probit analysis. Fitted psychometric functions have two parameters; mean (μ : the interpolated 50% crossover point) and standard deviation (σ : represents the slope of the psychometric function). The μ represents the point of subjective equivalence (PSE) and adaptation-effect degree, and the σ represents the precision of judgment. We then calculated the average μ and σ of two functions ("high" proportion and "high" and "same" proportion) in order to examine the PSE.

The μ and σ were calculated in each condition for all participants and separately analyzed by a 2 (Group: AWS vs. ANS) \times 3 (Adapted FC: -4 vs. ± 0 vs. $+4$) mixed-design analysis of variance with Shaffer's multiple comparisons. The μ in the adapted condition (FC: -4 or $+4$) denotes the susceptibility of disruption by FAF. The σ denotes participant judgment consistency. We regarded the ± 0 condition as the control condition for comparisons as there should be no adaptation effect.

In addition, to examine the relationship between the adaptation effect (i.e., $\Delta\mu$) and consistency of judgment (i.e., σ), Pearson's correlation analysis was performed between the $\Delta\mu$ (difference between adapted condition -4 or $+4$ and control condition ± 0) and σ within each group. The difference of μ between adapted and control conditions denotes the degree of the adaptation effect (i.e., a small difference between e.g., the -4 and ± 0 conditions denotes that adaptation was minor and thus that adaptation of FAF had limited influence, and vice versa). We used σ but not $\Delta\sigma$, because we assumed that the parameter that frequency adaptation changes was their perception of voice pitch (upward or downward; μ).

3. Results

We show the results of μ , σ , and the correlation between $\Delta\mu$ and σ of the calculated psychometric functions in Figs. 2, 3, and 4.

The μ (Fig. 2) showed a significant main effect of condition ($F(2, 76) = 117.70$, $p < 0.001$, $\eta_G^2 = 0.580$) but not of group ($F(1, 38) = 0.51$, $p = 0.480$, $\eta_G^2 = 0.007$), and an interaction between FC and group ($F(2, 76) = 4.13$, $p = 0.020$, $\eta_G^2 = 0.046$). In post-hoc analysis, in both groups, ± 0 was significantly higher than the -4 condition (AWS: $t(19) = 9.63$, $p < 0.001$, $d = 2.08$; ANS: $t(19) = 7.92$, $p <$

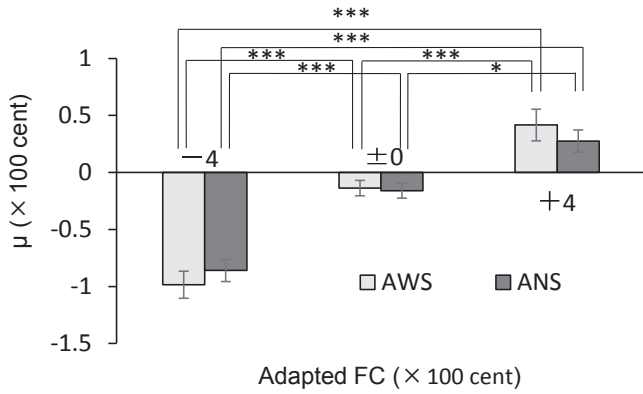


Fig. 2 Comparison of the mean μ of the psychometric function. Error bars represent standard errors (SEs). *** $p < 0.001$, * $p < 0.05$.

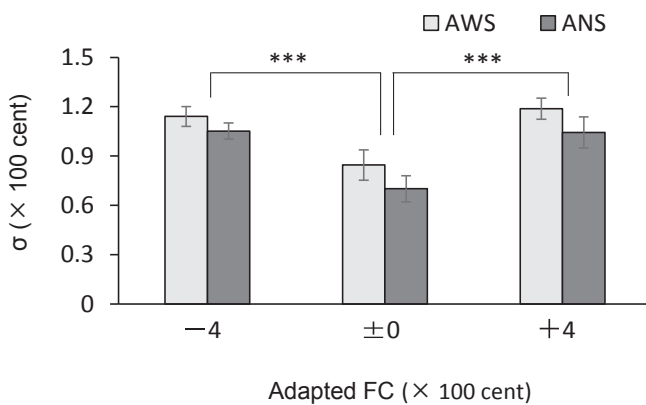


Fig. 3 Comparison of the mean σ of the psychometric function. Error bars represent standard errors (SEs). *** $p < 0.001$.

0.001, $d = 2.00$), +4 was significantly higher than the ± 0 condition (AWS: $t(19) = 4.44$, $p < 0.001$, $d = 1.20$; ANS: $t(19) = 2.75$, $p = 0.013$, $d = 0.73$), and +4 was significantly higher than the -4 condition (AWS: $t(19) = 9.84$, $p < 0.001$,

$d = 2.58$; ANS: $t(19) = 8.87$, $p < 0.001$, $d = 2.31$). However, no difference was found between groups in any of the FC conditions (-4: $F(1, 38) = 0.74$, $p = 0.395$, $\eta_G^2 = 0.019$, 0: $F(1, 38) = 0.06$, $p = 0.796$, $\eta_G^2 = 0.002$, +4: $F(1, 38) = 3.87$, $p = 0.056$, $\eta_G^2 = 0.092$).

The σ (Fig. 3) showed a significant main effect of condition ($F(2, 76) = 51.63$, $p < 0.001$, $\eta_G^2 = 0.203$) but no significant main effect of group ($F(1, 38) = 1.90$, $p = 0.18$, $\eta_G^2 = 0.04$) nor an interaction between them ($F(2, 76) = 0.36$, $p = 0.70$, $\eta_G^2 = 0.001$). This signifies that there was no difference in prediction precision in the group comparison. In post-hoc analysis, the σ in the 0 condition was significantly higher than that in the -4 condition ($t(38) = 8.83$, $p < 0.001$, $d = 1.05$) and in the +4 condition ($t(38) = 8.17$, $p < 0.001$, $d = 0.96$) in the within-group comparison.

Analysis of the correlation coefficient between $\Delta\mu$ and σ showed that while there was no significant correlation between $\Delta\mu$ and σ ($r = 0.025 \sim 0.292$, $p = 0.212 \sim 0.917$) in the ANS group, there was a significant correlation in the AWS group ($r = 0.446 \sim 0.495$, $p = 0.027 \sim 0.049$) (Fig. 4 as an example among the combinations of $\Delta\mu$ and σ) (Pearson's correlation analysis, $n = 20$ in the AWS and ANS group).

4. Discussion

Our results were partly consistent with a previous study using a DAF adaptation experiment [4]. We found a significant positive correlation between $\Delta\mu$ and σ in the AWS group, but not in the ANS group (Fig. 4). This suggests that the reliability of internal prediction differed within AWS, and AWS with reduced internal reliability appeared to compensate by relying greatly on auditory feedback information [4].

However, we did not find a group difference of μ in any of the adapted FC conditions (-4 or +4). This result is in marked contrast with a DAF adaption [4] and suggests that there is no difference in reliance on auditory feedback of frequency-information between AWS and ANS. This, in turn, suggests that the sensorimotor deficit in AWS may be timing-specific but may not be directly related to frequency (Fo) processing. This idea is consistent with the proposal that the internal prediction of sensory feedback after a motor

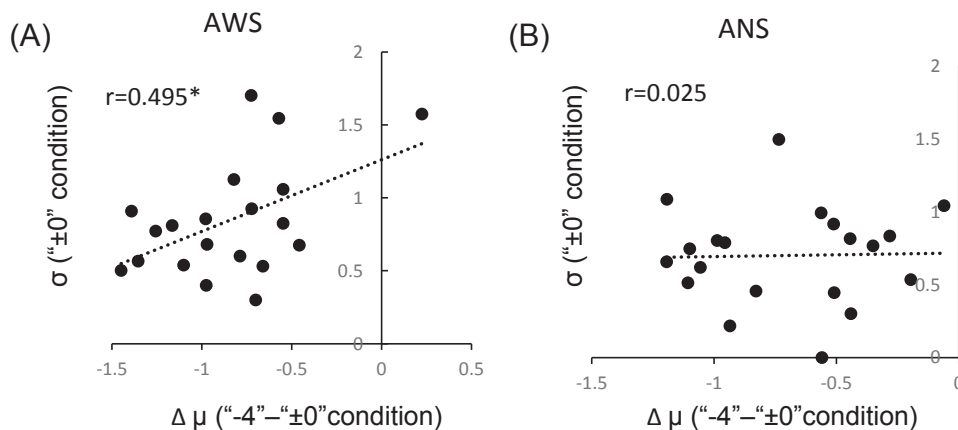


Fig. 4 Pearson's correlation between the $\Delta\mu$ and σ . (A) Distribution of AWS (B) Distribution of ANS. ** $p < 0.01$.

command could be separated into timing and other information [9]. Thus the results of our present and previous studies [4] could be reconciled if we assume that the brain separately encodes timing and frequency (Fo) information.

Stuttering is a disorder involving abnormal timing-processing [11] and this is supported by the fact that some AWS enhance their fluency under DAF conditions [12]. Our finding could add to the literature that stuttering is associated with sensorimotor processing deficits of timing information.

A further implication of the findings is that there is no group difference in μ . The sample size was relatively small but the effect size was medium to high in the present study. Thus, we are not willing to conclude that there is no difference and this should be addressed in future studies. We also did not take into consideration delays of a few milliseconds that could occur when participants' voices were fed back through the auditory effector.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas: Constructive Developmental Science (25119503) from MEXT and KAKENHI (19J10340, 19H00634, 19H05493) from JSPS.

References

- [1] B. Guitar, *Stuttering: An Integrated Approach to Its Nature and Treatment*, 5th ed. (Lippincott Williams & Wilkins, Baltimore, MD, 2018).
- [2] O. Civier, D. Bullock, L. Max and F. H. Guenther, "Computational modeling of stuttering caused by impairments in a basal ganglia thalamo-cortical circuit involved in syllable selection and initiation," *Brain Lang.*, **126**, 263–278 (2013).
- [3] G. Hickok, J. Houde and F. Rong, "Sensorimotor integration in speech processing: Computational basis and neural organization," *Neuron*, **69**, 407–422 (2011).
- [4] D. Iimura, N. Asakura, T. Sasaoka and T. Inui, "Abnormal sensorimotor integration in adults who stutter: A behavioral study by adaptation of delayed auditory feedback," *Front. Psychol.*, **10**, 2440 (2019).
- [5] L. Max, F. H. Guenther, V. L. Gracco, S. S. Ghoss and M. E. Wallace, "Unstable or insufficiently activated internal models and feedback-biased motor control as source of disfluency: A theoretical model of stuttering," *Contemp. Issues Commun. Sci. Disord.*, **31**, 105–122 (2004).
- [6] L. Max and A. Daliri, "Limited pre-speech auditory modulation in individuals who stutter: Data and hypotheses," *J. Speech Lang. Hear. Res.*, **62**(8S), 3071–3084 (2019).
- [7] S. Brown, R. J. Ingham, J. C. Ingham, A. R. Laird and P. T. Fox, "Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies," *Hum. Brain Mapp.*, **25**, 105–117 (2005).
- [8] D. M. Wolpert, J. Diedrichsen and J. R. Flanagan, "Principles of sensorimotor learning," *Nat. Rev. Neurosci.*, **12**, 739–751 (2011).
- [9] M. Tsakiris, P. Haggard, H. Franck, N. Mainy and A. Sirigu, "A specific role for efferent information in self-recognition," *Cognition*, **96**, 215–231 (2005).
- [10] M. Lincoln, A. Packman and M. Onslow, "Altered auditory feedback and the treatment of stuttering: A review," *J. Fluency Disord.*, **31**, 71–89 (2006).
- [11] P. A. Alm, "Stuttering and the basal ganglia circuits: A critical review of possible relations," *J. Commun. Disord.*, **37**, 325–369 (2004).
- [12] J. Kalinowski, A. Stuart, S. Sark and J. Armson, "Stuttering amelioration at various auditory feedback delays and speech rates," *Int. J. Lang. Commun. Disord.*, **31**, 259–269 (1996).