Title	Neural basis underlying vocal and speech control using auditory feedback
Sub Title	
Author	増田, 早哉子(Masuda, Sayako) 酒井, 奈緒美(Sakai, Naomi) 森, 浩一(Mori, Koichi)
Publisher	Centre for Advanced Research on Logic and Sensibility The Global Centers of Excellence Program, Keio University
Publication year	2009
Jtitle	CARLS series of advanced study of logic and sensibility Vol.2, (2008.) ,p.115- 125
JaLC DOI	
Abstract	
Notes	Part 1: Brain and Evolution
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO12002001-20090331-0115

11

Neural Basis Underlying Vocal and Speech Control Using Auditory Feedback

Sayako Masuda[1,2,3], *Naomi Sakai*[3,4], *and Koichi Mori*[3]

[1] Center for Advanced Research on Logic and Sensibility (CARLS), Keio University
[2] Nihon University
[3] National Rehabilitation Center for People with Disabilities
[4] Mejiro University

I. Introduction

Speech is one of the most important tools for communication in humans. In this chapter, we present studies investigating the neural basis for the control of speech using altered auditory feedback techniques. Auditory feedback, in which one's own voice or speech is monitored in real time, plays a critical role in normal speech production or vocalization in many vertebrates including humans, bats, and songbirds (Smotherman, Zang, and Metzner, 2003; Smotherman, 2007). Vocal pitch and intensity are either involuntarily or voluntarily controlled with auditory feedback while speaking. The Lombard effect (Lane and Tranel, 1971) is a phenomenon in which a speaker involuntarily raises his or her vocal intensity and modifies articulation in response to a high level of sound in the background. The development of normal articulation of speech production also depends on auditory feedback (Ando, Yoshino, Shimizu, and Itabashi, 1999).

Altered auditory feedback (AAF) has been introduced to study the effects of auditory feedback in speech and thereby elucidating the physiological and pathological speech control mechanisms. Masking auditory feedback (MAF), frequency-altered feedback (FAF) and delayed auditory feedback (DAF) are some forms of speech alteration used for research (Lincoln, Packman, Onslow, 2006). This chapter presents research examining effects on behavioral responses and neurological mechanisms using two kinds of altered auditory feedback methods; delayed auditory feedback (DAF) and transformed auditory feedback (TAF). First, we review the behavioral studies, and then the neural studies using these AAF techniques.

II. Altered Auditory Feedback and Behavioral Performance

1. Delayed Auditory Feedback (DAF)

Delayed auditory feedback is a technique in which the speaker's verbal output is fed back to his/her ears after a short delay, typically 50–200 milliseconds. Lee (1950) first reported the disrupting effects of DAF on speech of normally fluent speakers. Use of the DAF technique can cause stuttering, speech production to slow, or the speaker to raise their voice in pitch and/or volume; or to stop speaking completely. However, later studies do show that DAF may improve speech fluency of person who stutter (Lawrence and Barclay III, 1998; Van Borsel, Reunes, and Van den Bergh, 2003; Stuart, Kalinowski, Rastatter, Saltuklaroglu, and Dayalu, 2004; Van Borsel, Sierens, and Pereira, 2007). These data provide evidence that DAF is a potentially helpful technique in the treatment of stuttering (Stuart, Kalinowski, and Rastatter, 1997; Sparks, Grant, Millay, Walker-Baston, and Hynan, 2002; Stuart and Kalinowski, 2004; Sakai, Mori, Ozawa, and Mochida, 2008). Several devices with DAF are commercially available for the person who stutter, ranging from desk-top models to battery-powered devices worn in the ear canal. For example, Sakai and her colleague (2008) demonstrated the effectiveness of a behind-the-ear DAF device in reducing stuttering in the daily life of adults who stutter.

2. Transformed Auditory Feedback Method (TAF)

Although DAF has been used intensively in speech and disfluency research, it is difficult to analyze its effect by relating the delay time to its effect size quantitatively, because DAF often disrupts speech. In order to understand the finer interaction between speech production and perception under auditory feedback, it is desirable to use a small perturbation which does not totally disrupt speech but introduces detectable effects in produced speech (Kawahara, 1993a; 1993b; 1994). Kawahara (1993a) developed a transformed auditory feedback method (TAF) for quantitatively measuring the effect of auditory feedback onto the voice pitch (fundamental frequency, F0). His TAF method adds subliminal, fast parametric perturbations in the fundamental frequency (F0) of the feedback voice that would not disrupt vocalization but allow detection of possible perception/production interactions. The perturbation is inserted into the feedback loop of F0 according to a maximum length sequence (M-sequence). The interaction between speech production and perception under auditory feedback can be extracted by correlation analysis of the perturbation and the output vocal F0 changes, owing to the nature of M-sequence. Kawahara (1993a; 1993b; 1994) revealed with TAF that the produced speech shows a compensatory response to the perturbation with a 150 millisecond peak response latency on average. This method provides a non-invasive measurement method of auditory processes using one’s own voice as a probe.

Sato and his colleague (2002) investigated the voice control characteristics of persons who stutter using the TAF method. Their data reveal that the cross-correlation functions between the F0's of auditory feedback and uttered voice for some of these persons differed from those of people who do not stutter. Thus, abnormal auditory feedback loops may exist in some of the persons who stutter; suggesting that the TAF method may be useful in distinguishing subgroups among persons who stutter (Sato et al., 2002).

III. Auditory Feedback and Neural Mechanisms of Speech Control

1. Delayed Auditory Feedback (DAF)

Hirano and his colleague (1997) research indicates that increased activation occurs in the superior temporal gyrus (STG) as a function of DAF for persons who do not stutter (Hirano, Kojima, Naito, Honjo, Kamoto, Okazawa, Ishizu, Yonekura, Nagahama, Fukuyama, Konishi. 1997). On the other hand, Hashimoto and Sakai (2003) provide data showing increased activation in the bilateral temporo-parietal regions in addition to the STG under the DAF condition in persons who do not stutter. Watkins and his colleague (2008) performed fMRI experiments comparing activation locations among persons who stutter and those who do not stutter during overt sentence reading under normal feedback, DAF conditions. Their results show increased activation in the STG and the right inferior frontal cortex under the DAF condition relative to the normal feedback conditions in both groups. (Watkins, Smith, Davis, Howell, 2008)

2. Transformed Auditory Feedback (TAF)

There are two fMRI studies using a variant of the TAF experimental method described in Kawahara (1993a). Toyomura and his colleague (2007), use a method in which subjects have to maintain their normal voice pitch in the presence of distinctly inserted pitch shifts occurring two or three times during each trial. Their results comparing the variant-TAF against the non transformed AF conditions reveal significant activations in the supramarginal gyrus, the prefrontal area, the anterior insula, the superior temporal area and the intraparietal sulcus in the right hemisphere; but only the premotor area in the left hemisphere during the variant-TAF experimental condition (Toyomura, Koyama, Miyamaoto, Terao, Omori, Murohashi, Kuriki, 2007). Tourville, Reilly, and Guenther, (2008) used an altered auditory feedback method in which the first formant frequency of their speech is unexpectedly shifted. Their results show a significant contrast between pitch-shift AF and no-shift AF conditions, revealing significant activations in the bilateral superior temporal gyrus and left middle temporal gyrus.

There are some differences between the original TAF experimental condition (e.g., Kawahara, 1993a) and these variant TAF condition, such as the magnitude of pitch shift, subject's awareness, voluntariness, and times of pitch shift.

IV. Current Research Presented

We are investigating the neural mechanisms underlying speech production and its feedback control using functional magnetic resonance imaging (fMRI). In One study we use the DAF technique to examine cortical activation patterns in both persons who stutter and for persons who do not stutter (Sakai, Masuda, Shimotomai, Mori, in press). The second study uses the TAF technique with persons who do not stutter (Masuda, Sakai, Okazaki, Kamatani, Mori, Nakajima, in preparation). Both studies were approved by the Ethics Committee of National Rehabilitation Center for Persons with Disabilities.

We obtain functional data using an event-triggered sparse sampling technique (Yang, Engelien, Engelien, Xu, Stern, and Silbersweig, 2000; Le, Patel, and Roberts, 2001; Engelien, Yang, Engelien, Zonana, Stern, and Silbersweig, 2002; Fu, Vythelingum, Brammer, Williams, Amaro, Andrew, , Yaguez, van Haren, Matsumoto, and McGuire, 2006). Single volume image data are obtained at a certain time delay after each trial onset. The delay period is inserted to obtain the blood-oxygen-level-dependent (BOLD) fMRI data at or near the peak of the hemodynamic response to vocalization (estimated to occur at approximately 4–7 s after speech or vocalization). The procedure continues with the next trial starting following another delay period.

The sparse sampling design affords several important advantages. First, it allows subjects to speak in relative silence, or in a more natural speaking condition, rather than speaking during the loud scanner noise. Second, it permits online digital signal processing to be applied to the speech signal for adding the perturbation (DAF or TAF), which is not possible in the presence of scanner noise because any feedback voice, at a reasonable sound level, would be masked by the noise. Finally, since scanning is carried out only after speech ceases, it eliminates artifacts due to movement of the head and changing volume of the oral cavity during speech.

1. Delayed Auditory Feedback Effects Represented in the Brain

1.1. Methods

Right handed Native Japanese speakers who do not stutter and who stutter participated in the study. They were instructed to read aloud 100 Japanese sentences, the half of which were to be read under a non-altered feedback (NAF) condition, and the other half under a DAF condition. The delay in the feedback was fixed at 200 ms for subjects who did not stutter and optimized for subjects who do stutter for fluency (50 to 20 ms). Viewing a plus sign without utterance served as a control condition. The three conditions (Control, NAF, and DAF) were interleaved serially. The visual sentence stimulus was presented for 4s at each trial under the NAF and DAF conditions, and subjects were instructed to read aloud only during the presentation. Spatial preprocessing and statistical analyses of the brain images were performed using SPM99.

1.2. Results and Discussion

Higher activations were observed in the DAF condition as compared to the activation levels under the NAF condition for subjects who did not stutter. These differences in activation levels were observed in the right and left superior temporal gyrus, the right middle temporal gyrus, and the right inferior frontal gyrus (BA45) ($p < 0.001$, uncorrected).

On the other hand, higher activations occurred under the DAF condition for speakers that stutter, as compared to activation levels under the NAF condition. These differences were observed in the right superior temporal sulcus and superior temporal gyrus, the right and left middle temporal gyrus, the left supplementary motor area and the right middle frontal gyrus (premotor area) ($p < 0.001$, uncorrected).

Our results reconfirmed most of the previous findings with slight differences on the detailed activation areas. Some of the differences may be attributable to different sampling biases of participants between research institutions, especially that of the distribution of stuttering severity, and to the details of the methods used, which would require further large scale studies.

2. Transformed Auditory Feedback Effects Represented in the Brain

The TAF method we employed dealt with the F0 (vocal pitch) control by auditory feedback. In order to investigate the relevant cortical activations to F0 control, we performed fMRI experiments during vocalization under the TAF condition for speakers that do not stutter (Masuda et al., 2008).

2.1 Methods

Neurologically normal subjects who did not stutter participated in the study. They were all right-handed, native Japanese speakers. None had been trained as a singer. They were instructed to repeatedly vocalize the vowel /a/ for at least 5 seconds at a visual cue in the MRI gantry. They were instructed to keep the voice pitch as stable as possible. TAF and non-transformed auditory feedback (NAF) at a matching level as TAF of the subject's own voice were heard in a pseudo-random order through a pair of condenser earphones. Each subject repeated the sustained vocalization 10 times per session, and went through 4 sessions. Spatial preprocessing and statistical analyses of the acquired brain images were performed using SPM5.

2.2. Results and Discussion

Comparison of the TAF condition against the NAF condition revealed activated areas in the right MTG, the left STG, the bilateral precentral gyrus and the right posterior insula ($p < 0.0005$, uncorrected, $k >= 50$). We conclude that these regions play important roles in the pitch control by auditory feedback.

V. Discussion

In these two studies, we performed the fMRI experiments in order to investigate the neural mechanisms of speech and/or vocalization control under the AAF conditions. Several areas within the auditory cortex were found activated by both our DAF and TAF conditions. Previous mammal and human studies also reported that the auditory cortex is sensitive to auditory feedback during vocal production (McGuire et al., 1996; Eliades and Wang, 2008). In previous DAF and TAF studies, the same posterior superior temporal area is activated (Hirano et al., 1997; Hashimoto and Sakai, 2003; Toyomura et al. 2007; Tourville et al., 2008).

A significant activation in the right middle temporal gyrus was also commonly observed in the present DAF and TAF studies. One such study using a variant TAF also shows significant activation in the middle temporal gyrus (Tourville et al., 2008) and supports the finding of the present research.

In the DAF study, the non stuttering speakers showed higher activation levels in the right supplementary motor area and superior temporal gyrus when compared to the activation levels observed with persons who stutter. This finding was under both of NAF and DAF conditions. These observations are in contrast to the observed activation sites associated with the subject group of speakers who stutter; whereas, the speakers who do not stutter showed the higher activation in the inferior frontal gyrus (the right BA45; the bilateral BA46) under the DAF condition than in the NAF condition respectively.

The involvement of the inferior frontal gyrus, especially on the right side, and the superior temporal gyrus is a common observation related to impaired fluency (Fox, Ingham, Ingham, Zamarripa, Xiong, Lancaster, 2000; Watkins et al., 2008). Activation levels in these right inferior frontal gyrus and bilateral superior temporal areas of persons who stutter tends to be smaller than observed for people who do not stutter.

In the TAF study, the neural substrate for the involuntary, subliminal vocal pitch control using online (not delayed) auditory feedback was investigated. The results show that the neural network subserving the vocal pitch control was distributed over the middle temporal gyrus, superior temporal gyrus, precentral gyrus and the right insula. In this study, the precentral gyrus and the right insula, uniquely activated in this condition, may concern motor control for pitch compensation.

Reference

Ando, T., Yoshino, T., Shimizu, Y., Itabashi, Y., 1999, Relations Among Articulation Scores in Speech Perception, Articulation Scores in Speech Production, and Hearing Level in Children with Hearing Impairments, The Japanese Journal of Special Education, 36, 49-57.

Chase, R. A., Sutton, S., First, D., Zubin, J., 1961, A developmental study of changes in behavior under delayed auditory feedback. Journal of Genetic Psychology, 99, 101-112.

Degermana, A., Rinnea, T., Salmia, J., Salonen, O., Alho, K., 2006, Selective attention to sound location or pitch studied with fMRI. Brain Research, 1077, 123-134.

Eliades, S. J., Wang, X., 2005, Dynamics of auditory-vocal interaction in monkey auditory cortex. Cereb Cortex, 15, 1510–23.

Engelien, A., Yang, Y., Engelien, W., Zonana, J., Stern, E., Silbersweig, D.A., 2002. Physiological mapping of human auditory cortices with a silent event-related fMRI technique. NeuroImage 16, 944–953.

Fox, P. T., Ingham, R. J., Ingham, J. C., Zamarripa, F., Xiong, J. H., Lancaster, J. L., 2000, Brain correlates of stuttering and syllable production: A PET performance-correlation analysis. Brain. 123, 1983-1984.

Fu, C. H. Y., Vythelingum, G. N., Brammer, M. J., Williams, S. C. R., Amaro, E. Jr., Andrew, C. M., Yaguez, L., van Haren, N. E. M.,Matsumoto, K., & McGuire, P. K., 2006. An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. Cerebral Cortex, 16, 969–977.

Hashimoto Y, Sakai K. L., 2003, Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. Hum Brain Mapping, 20, 22-28.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., Konishi, J. 1997, Cortical processing mechanism for vocalization with auditory verbal feedback. Neuroreport. 7, 2379-2382.

Kawahara, H. 1993a, Implications of Transformed Auditory Feedback in Hearing Research. ITEJ Technical Report Vol.17, No.53, 1-6.

Kawahara, H., Williams, J. C., 1993b, The Institute of Electronics, Information and Communication Engineers Bibliography Analysis of Pitch Perturbation Effects by Transformed Auditory Feedback. IEICE technical report. Speech Vol.93, No.155.

Kawahara, H., 1994, Interactions between Speech Production and Perseption under Transformed Auditory Feedback. Technical report of IEICE. HC Vol.93, 33-40

Lane, H., Tranel, R., 1971, The lombard sign and the role of hearing in speech. Journal of Speech and Hearing Research, 14, 677–709

Lawrence, M., Barclay III, D. M., 1998, Stuttering: A brief review, American Family Physician, 57, 2175-2180.

Le, T. H., Patel, S., Roberts, T. P., 2001. Functional MRI of human auditory cortex using block and event-related designs. Magn. Reson. Med. 45, 254–260.

Lee, B. S., 1950, Some effects of side-tone delay. Journal of Acoustic Society America, 22, 824–826

Lincoln M, Packman A, Onslow M. 2006, Altered auditory feedback and the treatment of stuttering: A review. Journal of Fluency Disorders, 31, 71-89.

Macleod, J., Kalinowski, J., Stuart, A., Armson, J., 1995, Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. Journal of Communication Disorders, 28, 217–228.

Masuda, S., Okazaki, S., Sakai, N., Kamatani, D., Mori, K., & Nakajima, Y. (2008) The neural basis subserving pitch control revealed by the transformed auditory feedback: A magnetic resonance imaging study. NEUROSCIENCE 2008 (Washington Convention Center, Washington D.C., U.S.A, 11/15-19)

Masuda, S., Sakai, N, Okazaki, S., Kamatani, D., Mori, K., Nakajima, Y., Neural basis underlying vocal and speech production using transformed auditory feedback: An fMRI study. (in preparation)

McGuire, P. K., Silbersweig, D. A., Frith, C. D., 1996, Functional neuroanatomy of verbal self-monitoring. Brain 119, 907-917.

Sakai, N., Mori, K., Ozawa, E., & Mochida, A. 2008, Effectiveness of a behind-the-ear device delivering delayed auditory feedback on stuttering in the daily life of an adult stutterer. The Japan Journal of Logopedics and Phoniatrics, 49, 107–114.

Sakai, N., Masuda, S., Shimotomai, T., Mori, K., Brain activation in adults who stutter under delayed auditory feedback: An fMRI study. International Journal of Speech–Language Pathology (in press)

Sato, Y., Mori, K., Fukushima Y., 2002 Temporal Characteristics of Fundamental Frequency Control by Auditory Feedback and its Application to Stuttering Evaluation. IEICE technical report, 101, 25-30.

Smotherman M. S., 2007, Sensory feedback control of mammalian vocalizations. Behavior Brain Research. 4, 315-326.

Smotherman M, Zhang S, Metzner W., 2003, A neural basis for auditory feedback control of vocal pitch. Journal of Neuroscience, 15, 1464-1477.

Soderberg, G. A., 1969, Delayed auditory feedback and the speech of stutterers. Journal of Speech and Hearing. 4, 20-29.

Sparks, G., Grant, D., Millay, K., Walker-Baston, D., Hynan, L., 2002, The effect of fast speech rate on stuttering frequency during delayed auditory feedback. Journal of Fluency Disorders, 27, 187–201.

Stuart, A., Kalinowski, J. 2004, The perception of speech naturalness of post-therapeutic and altered auditory feedback speech of adults with mild and severe stuttering. Folia Phoniatrica et Logopaedica, 56, 347–357.

Stuart, A., Kalinowski, J., Rastatter, M., Saltuklaroglu, T., & Dayalu, V. 2004, Investigations of the impact of altered auditory feedback in-the-ear devices on the speech of people who stutter: Initial fitting and 4-month follow-up. International Journal of Language and Communication Disorders, 39, 93–119.

Stuart, A., Kalinowski, J., Rastatter,M. P., 1997, Effect of monaural and binaural altered auditory feedback on stuttering frequency. Journal of the Acoustical Society of America, 101, 3806–3809.

Tourville, J. A., Reilly, K. J., Guenther, F. H., 2008, Neural mechanisms underlying auditory feedback control of speech. Neuroimage, 1, 1429-1443.

Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., Kuriki, S., 2007, Neural correlates of auditory feedback control in human. Neuroscience, 11, 499-503.

Van Borsel, J., Reunes, G., & Van den Bergh, N. 2003, Delayed auditory feedback in the treatment of stuttering: Clients as consumers. International Journal of Language and Communication Disorders, 38, 119–129.

Van Borsel, J., Sierens, S., Pereira, M. M. 2007, Using delayed auditory feedback in the treatment of stuttering: evidence to consider. Pro-Fono Revista de Atualizacao Cientifica, 19, 323-332.

Watkins, K. E., Smith, S. M., Davis, S., Howell, P. 2008, Structural and functional abnormalities of the motor system in developmental stuttering. Brain, 131, 50-59.

Yang, Y., Engelien, A., Engelien, W., Xu, S., Stern, E., Silbersweig, D.A., 2000. A silent event-related functional MRI technique for brain activation studies without interference of scanner acoustic noise. Magn. Reson. Med. 43, 185–190.