

Vocal Adaptation in Perturbed Auditory Feedback for L2 Learners of Mandarin^{*}

Li-Hsin NING
Department of English
National Taiwan Normal University

Abstract

This research explores the vocal adaptation in L2 learners of Mandarin by constructing a pitch-shift paradigm that incorporates adaptive changes (i.e., continuously increasing or decreasing changes) in pitch in auditory feedback during vocalization. This study examined the productions of sustained /a/ and /ma1/ (“mother” in Mandarin) in L2 beginners, L2 advanced learners, and native Mandarin speakers. The results show that L2 advanced learners and native Mandarin speakers, but not L2 beginners, had reduced reliance on auditory feedback to monitor vocal pitch. Sensorimotor adaptation was observed for all the groups, suggesting that the online recalibration of the internal model is automatic. The online recalibration is also task-dependent, depending on the requirement of vocal accuracy and the closeness of intended pitch and perturbed pitch.

Keywords: sensorimotor adaptation, pitch-shift paradigm, internal model, L2 learners of Mandarin

^{*} This study was supported by the MOST Grant 105-2410-H-003-001. I thank Jian Leat Siah and Chien-min Kuo for assisting with data collection and participant recruitment.

1. Introduction

Sensory information such as auditory feedback along with motor (articulatory) command preparation has been theorized in the internal model within the nervous system. (Jordan and Rumelhart 1992; Wolpert et al. 1995; Perkell et al. 1997; Kawato 1999; Max et al. 2004; Guenther 2006; Lalazar and Vaadia 2008). A sub-system called the *inverse internal model* takes the desired movement goal and prepares the motor commands required to accomplish it. An efference copy generated by the motor control system contains a copy of the expected motor commands that is sent to sensory processing regions. Another sub-system called the *forward internal model* predicts the sensory state resulting from motor commands (afferent signals), which is essentially the predicted outcome of a movement. A comparator then checks if the predicted outcome of the movement corresponds to the motor commands. If so, the motor commands are issued to the muscles. If there is a mismatch, then an error signal is generated to modify the motor commands. Therefore, sensory feedback used in control of action (including speech) is considered as a closed-loop system in the nervous system, which monitors intended movements and reduces errors if any. The internal model is conceived as a malleable neural system that predicts the relationships between past, current and future states of the nervous system to provide movements that are task appropriate and simulate the relationships between motor commands, motor trajectories and sensory feedback (Guenther 1995; Kawato 1999; Callan et al. 2004; Guenther 2006; Lalazar and Vaadia 2008).

Auditory feedback is particularly important for learning to sing or speak in that it is associated with checking vocal output against a copy of the expected articulatory movement. The frequency-altered feedback (FAF) paradigm, where a short and artificial change in pitch (either upward or downward shift) is fed back to speakers during vocalizations, has been used to investigate the role of auditory feedback in vocal pitch error detection and correction. By comparing the pitch-altered feedback with a target pitch in the internal model established by memory or efference copy, speakers would get tricked into thinking they have made an error in the vocalization. In order to reduce the error, vocal responses to auditory perturbation (i.e., pitch-shift responses) are typically compensatory (Burnett et al. 1998; Larson et al. 2000; Hain et al. 2001; Burnett and Larson 2002; Chen et al. 2007; Liu et al. 2009). *Compensation*,

a reflex-like response, is the evidence of online control of speech through auditory-motor feedback loop.

Research has shown that singers and native Mandarin speakers have suppressed pitch-shift responses compared to nonsingers and non-native tonal speakers (Jones and Keough 2008; Ning et al. 2014; Ning et al. 2015). The *suppression* suggests that singers and native Mandarin speakers have robust, stable, and enhanced internal models for pitch, built through years of vocal or language training, which allows them to reduce the reliance on auditory feedback and to control their voice F0 through an internal reference (i.e., a strong feedforward command). Native Mandarin speakers also have shown compensatory pitch-shift responses with more rapid latencies than native English speakers, suggesting that the required ability to perceive and produce rapid changes in F0 in a tonal language would enable tone speakers to rapidly correct the mismatch between expected F0 and heard (perturbed) F0 (Burnett et al. 1998; Hain et al. 2000; Larson et al. 2001; Xu et al. 2004). Pitch-shift responses, therefore, could serve as a probe for examining the attainment of native-like tonal representations in the human brain.

This research further examines whether our speech motor commands would adapt to perturbed auditory feedback and whether the adaptation, like compensation and suppression, is subject to language experience. While suppression represents the robustness of internal model for pitch, *sensorimotor adaptation* may indicate the internal model's ability to recalibrate an internal reference online. Experimental evidence from pitch perturbation studies indicates that feedforward control of speech motion can be updated by continuous changes in auditory feedback. Instead of using constant pitch-shifts in auditory feedback during the entire experiment, adaptation studies use gradual but slight alterations in auditory feedback to modify speech output in adults. The adaptation experiments typically involve three phases: (i) the baseline phase where normal auditory feedback is fed to speakers; (ii) the learning phase where the pitch in auditory feedback is altered or increased/decreased successively; (iii) the test phase where speakers receive normal auditory feedback again. *Sensorimotor adaptation* could be observed in the learning phase, where the current voice F0 initiates at a similar F0 obtained on the previous trial. *Aftereffect* is examined by comparing the responses in the test phase with the responses in the baseline phase and

refers to the situation where speakers responded as if they heard perturbed auditory feedback in the test phase.

Previous adaptation studies mainly focused on the influence of vocal training experience. Singers can continuously update the internal models to accommodate the continuous changes in F0. However, nonsingers, lack of a precise internal representation of pitch, struggle with finding a correct pitch target so that the continuous changes in pitch would be seen as random errors instead of systematic changes (Jones and Keough 2008; Keough and Jones 2009). Research has shown inconsistent results on the aftereffect. While Jones and Keough (2008) argue that singers but not nonsingers showed the aftereffect, Keough and Jones (2009) and Keough et al. (2013) claim that the aftereffect appeared in both singers and nonsingers.

The sensorimotor adaptation and immediate aftereffect were also found in native Mandarin speakers' production of tone 1 and tone 2 (Jones and Munhall 2005). The findings indicate that speakers who have obtained an entrenched internal model for pitch are sensitive to continuously increasing or decreasing changes in pitch and thus would start remapping the auditory-motor relationship in a brief training session. While suppressed pitch-shift responses in constant pitch-altered feedback allow us to explore the robustness of internal model for pitch, sensorimotor adaptation and aftereffect open windows for observing the plasticity of internal model of pitch in language learners.

This study made an initial attempt to investigate the effects of language learning experience (native Mandarin vs. L2) on the remapping of auditory input and motor commands, using the FAF paradigm. For *compensation*, we investigate whether native Mandarin speakers would have suppressed responses in the gradual pitch-shifts as found in the constant pitch-shifts. For *sensorimotor adaptation*, we examine whether L2 learners, like native Mandarin speakers, can update the internal models to accommodate the continuous changes in F0. If so, it may suggest that the online recalibration is automatic. If not, it may suggest that volitional control could be involved in the online recalibration. For *aftereffect*, we are interested in the persistent effect of online recalibration. The presence or absence of aftereffect may reveal what the internal references look like before the experimental training is implemented.

2. Methods

2.1 Participants

Fourteen native speakers of Mandarin (8 females) and fourteen second language learners of Mandarin whose native languages were non-tonal languages participated in this study. Of the 14 second language learners, 7 (6 females) were beginners with two years of acquisition on average, and 7 (3 females) were advanced learners with approximately five years of acquisition. A self-report screening test and a hearing test with use of an MAICO pure tone audiometer (model MA 25) were administered to all participants. None of the participants reported a history of neurological or communication disorders. All participants passed a hearing test at 20 dB hearing level bilaterally at 125, 250, 500, 750, 1000, 2000, 3000, and 4000 Hz. All participants were right-handed and between the ages of 20-35. All participants received financial compensation for their time and informed consent was collected from each participant.

2.2 Apparatus

Participants were recorded with an AKG K240 headphone and a standalone microphone (Audio Tech ATR20) placed 2 cm away from the mouth. The speech signals were amplified with an ART Tube MP microphone preamplifier, sent to a Behringer firewire audio interface (model FCA610), and shifted in pitch with an Eventide Ultra-Harmonizer (model H7600). MIDI software (Max/MSP v.7 by Cycling 74) was used to control the timing, duration and magnitude of the pitch shifts via the Eventide Ultra-harmonizer. The participant's voice signal was amplified with a Samson S-phone Headphone Amplifier and played back to him/her with a gain of 10 dB SPL louder to reduce the possibility that speakers may hear their own nonshifted, bone-conducted signal. The participant's voice, altered feedback signal, and pitch-shift events (transistor transistor logic (TTL) pulses) were digitized at 5 kHz per channel by WinDag DI-720 (DATAQ Instruments), and recorded using WinDag Pro software (DATAQ Instruments).

2.3 Materials and Procedure

Participants attended four sessions of pitch production that differed in the target stimuli (/a/ with a comfortable steady pitch and /ma1/ with a high level tone in Mandarin) and the direction of pitch-shifts (upward shift and downward shift). The

order of the four sessions was counterbalanced across subjects. Each session consisted of 10 trials of baseline, 51 FAF trials, and 20 test trials (see Figure 1). During the baseline 10 trials, participants produced either /a/ or /ma1/ (hereafter, AH and MA1 for short) with normal auditory feedback. In the learning phase (i.e., the FAF trials), the auditory feedback was increased (up-shift) or decreased (down-shift) by 2 cents for each successive utterance until the feedback received was one semitone (100 cents) above or below the participant's pitch. The auditory feedback was shifted from the onset of each utterance until the end of the vocalization. There was no noticeable lag between the onset of the utterance and the onset of perturbed auditory feedback. In the test phase, participants produced 20 more utterances with normal auditory feedback. In each trial, a beep sound indicated the start of a trial, followed by a voice sample of the stimulus. After the voice sample, participants produced the same stimulus and held the vocalization for 2 seconds. Participants were told to produce the target stimuli with their comfortable steady pitch (i.e., the conversational pitch level) rather than imitate the sample's pitch. The intertrial interval was 1 second. The experiment lasted approximately 1 hour.

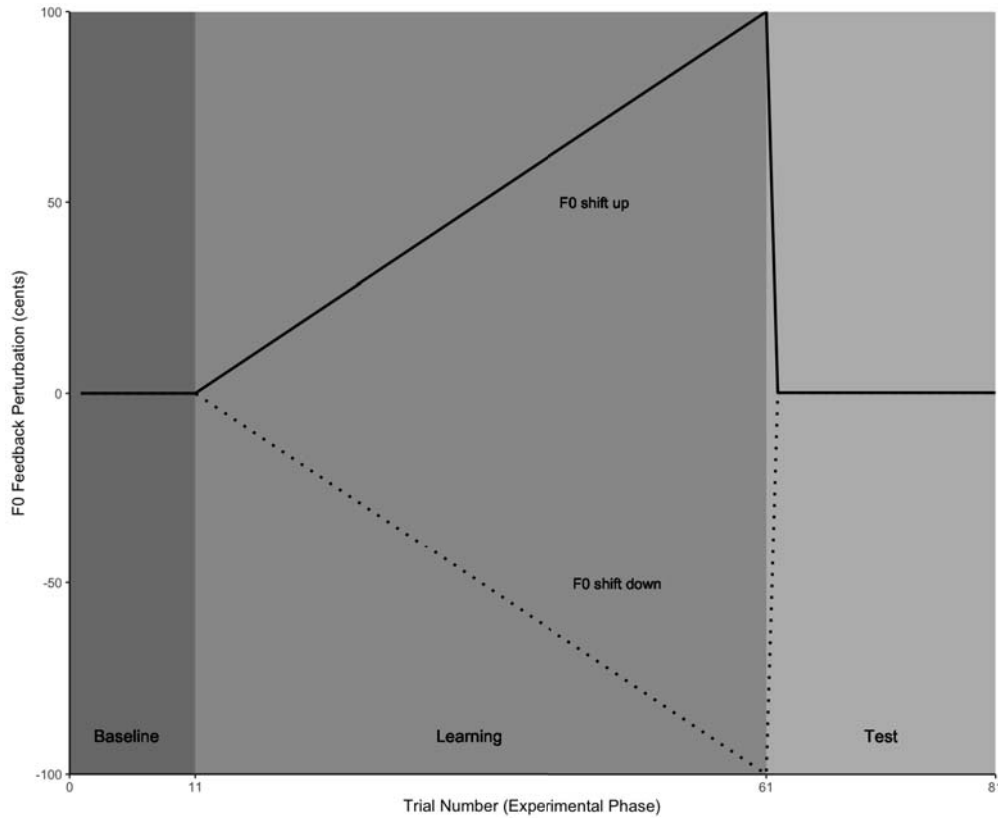


Figure 1: The Gradual Change in Frequency-altered Feedback (FAF)

During the baseline (trials 1-10), the pitch of auditory feedback was not altered. In the learning phase, auditory feedback was manipulated from trial 11 to trial 61, and shifted either upwards or downwards in 2 cent increments across trials until auditory feedback was 1 semitone (100 cents) above or below the participant's pitch. During the test trials (trials 62-81), auditory feedback was returned to normal.

2.4 Data Analysis

The raw signals in the pitch production were imported to MATLAB. The pitch values were tracked using the SWIPE pitch estimation algorithm (10 ms interval) (Camacho and Harris 2008; Scheerer and Jones 2012) then converted into cents (cents = $1200 \cdot \log_2 (F0/\text{baseline } F0)$, where the baseline F0 is the mean F0 in the baseline phase).

The trials in each session were subdivided into 8 blocks of ten trials: 1 block

of baseline trials (1-10), 5 blocks of FAF trials (11-21¹, 22-31, 32-41, 42-51, 52-61), and 2 blocks of test trials (62-71, 72-81). Two statistical analyses were conducted to determine the effects of the experimental factors on the *compensation and sensorimotor adaptation* (on the FAF trials) and *aftereffect* (when auditory feedback was returned to normal). To analyze the *compensation*, a repeated measures ANOVA with 3 (GROUP: L2 beginner, L2 Advanced learner, and native Mandarin speakers) x 5 (BLOCK: 11-21, 22-31, 32-41, 42-51, 52-61) x 2 (TASK: AH or MA1) x 2 (DIRECTION: upward and downward shift) factors was conducted on the FAF trials. Tukey's honestly significant difference (HSD) test was used for post-hoc analyses with an alpha level of .05. To analyze the *aftereffect*, t-tests were carried out on the baseline trials (1-10) and the first ten test trials (62-71). The first ten test trials were selected in order to (i) have an equal number of trials as in the baseline, and to (ii) reveal the initial responses when auditory feedback was returned to normal.

3. Results

3.1 Compensation and Sensorimotor Adaptation

The mean F0 values (cents) in the learning phase are presented in Figure 2. In order to evaluate the effect of DIRECTION, absolute pitch values were fitted into the ANOVA model. The ANOVA with 3 (GROUP: L2 beginner, L2 Advanced learner, and native Mandarin speakers) x 5 (BLOCK: 11-21, 22-31, 32-41, 42-51, 52-61) x 2 (TASK: AH or MA1) x 2 (DIRECTION: upward and downward shift) factors showed that there was a significant main effect of GROUP, $F(2,500) = 2.993$, $p < .05$. On average, L2 beginners had larger compensation (70 cents) than L2 advanced learners (61 cents) and native Mandarin speakers (56 cents), while there was no significant difference in compensation between L2 advanced learners and native Mandarin speakers.

The main effects of BLOCK, $F(4,500) = 24.373$, $p < .001$, and DIRECTION, $F(1,500) = 20.505$, $p < .001$, were also significant. Post-hoc tests showed that the compensation gradually increased from the first block of FAF trials to the fifth block

¹ There were 11 trials instead of 10 in the first block of FAF because the perturbation started from 0 cent. This classification ensures that each block ended at a multiple of 20 cents.

of FAF trials, but no significant differences were observed in the comparison between the second block and the third block, and the comparison between the fourth block and the fifth block (i.e., Trials 11-21 < Trials 22-31 = Trials 32-41 < Trials 42-51 = Trials 52-61). Since cents values were normalized to the mean F0 values of the baseline, the gradual F0 changes away from a cent value of 0 indicated sensorimotor adaptation where participants continuously updated the internal models for pitch. As for the main effect of DIRECTION, speakers had larger compensatory responses to downward shifts (71 cents) than to upward shifts (53 cents). There was no significant main effect of TASK, $F(1,500) = 2.073, p = .151$.

A significant interaction was found in GROUP x DIRECTION, $F(2,500) = 3.454, p < .05$. The simple main effect analyses showed that L2 beginners and native Mandarin speakers had larger compensation in the downward shift condition than in the upward shift condition. In the downward shift condition, L2 beginners (88 cents) demonstrated larger compensatory responses than L2 advanced learners (60 cents) and native Mandarin speakers (68 cents) (Bonferroni-adjusted p 's < .05). A significant interaction between GROUP and TASK was also observed, $F(2,500) = 8.751, p < .001$. The simple main effect analyses showed that L2 advanced learners and native Mandarin speakers had larger compensation in MA1 (L2: 66.911 cents; Mandarin: 66.225 cents) than AH (L2: 45.327 cents; Mandarin: 56.263 cents), while L2 beginners compensated more in AH (82.471 cents) than in MA1 (57.417 cents) (Bonferroni-adjusted p 's < .05). No other significant main effects or interactions were observed.

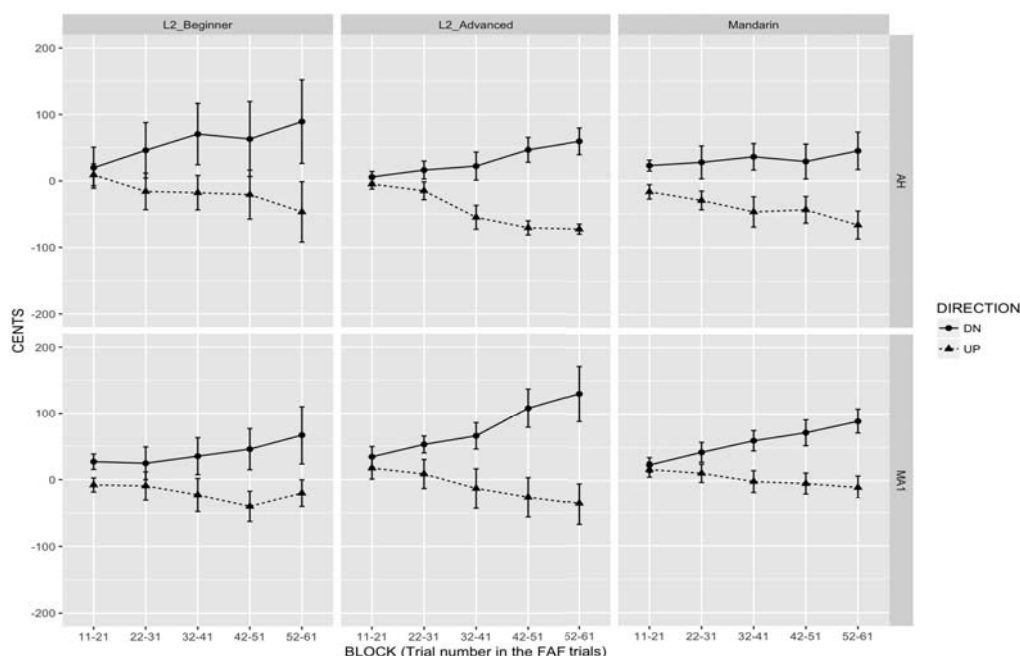


Figure 2: The Mean Pitch Values (Cents) across 5 Blocks (10 Trials Per Block) of FAF Trials

F0 was calculated relative to the mean pitch in the baseline for each condition. The auditory feedback was gradually (2 cent increments) shifted until the shifted pitch was 100 cents above (in the upward shift condition) or below (in the downward shift condition) the participant's pitch. Error bars represent the standard error of the mean.

3.2 Aftereffect

The mean F0 values (cents) in the baseline phase (1-10) and the first ten test trials (62-71) are presented in Figure 3. Corrected multiple t-tests were carried out on the baseline trials and the first ten test trials for each level of conditional factor. The purpose was to determine whether compensating for gradual FAF would result in aftereffects when auditory feedback was returned to normal. In the upward shift AH, L2 advanced learners' F0 values in the first ten test trials were significantly different from the F0 values in the baseline trials, $t(6) = 3.775$, $p < .01$. No other significant aftereffect was found in the production of AH.

In the downward shift MA1, all three groups demonstrated that the F0 values in the first ten test trials were significantly higher than the F0 values in the baseline (L2 beginners: $t(6) = -5.324, p < .01$; L2 advanced learners: $t(6) = -2.740, p < .05$; native Mandarin speakers: $t(6) = -2.674, p < .05$). This suggests that when auditory feedback was returned to normal, participants responded as if they had heard the altered feedback and thus showed the compensatory responses. No significant aftereffect was found in the upward shift MA1.

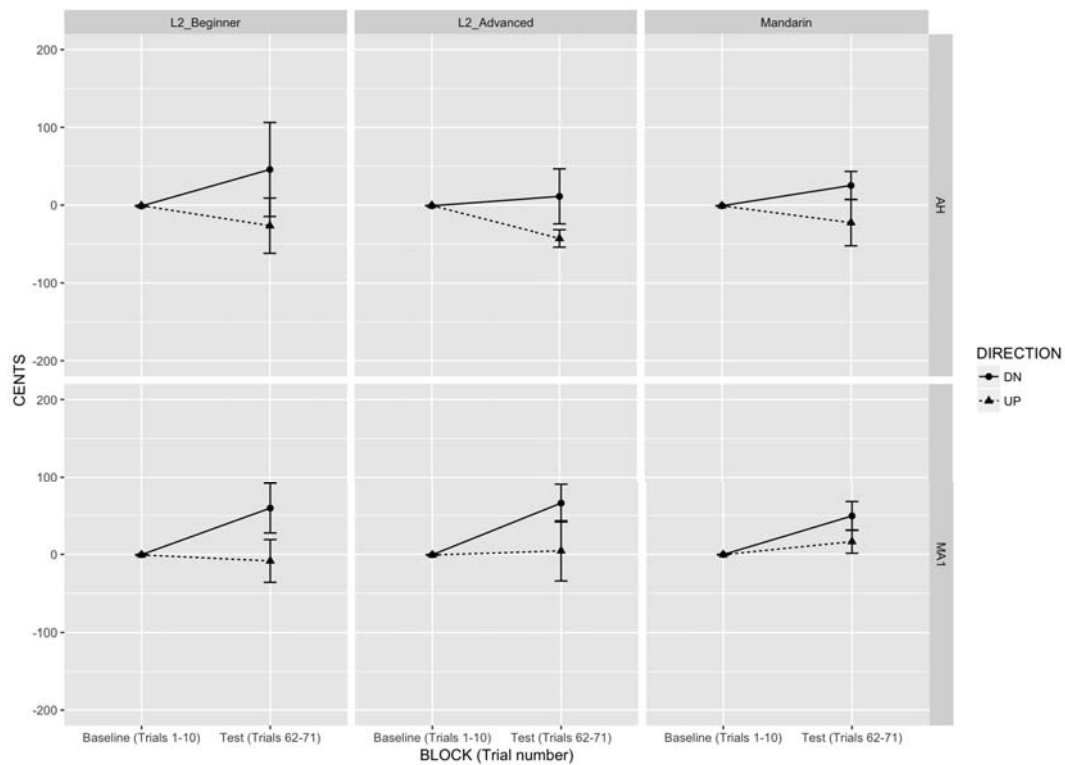


Figure 3: The Mean Pitch Values (Cents) in the Baseline Trials (1-10) and the First Ten Test Trials (62-71)

F0 was calculated relative to the mean pitch in the baseline for each condition. Participants received normal auditory feedback in both the baseline phase and the test phase. Error bars represent the standard error of the mean.

4. Discussion

The aim of this research was to examine whether remapping auditory-motor relationships appears in second language learners of Mandarin (beginners and advanced learners) when they are asked to vocalize the sustained vowel /a/ and the Mandarin syllable /ma1/ in the FAF paradigm. The results show that in general all the groups responded to the artificial pitch shifts in a direction opposite to the direction of pitch manipulation (i.e., compensatory responses). Compensation indicates that participants tried to maintain the intended pitch while receiving frequency perturbation in auditory feedback. The magnitude of compensation increased gradually in the FAF trials as the absolute magnitude of the pitch manipulation increased. This suggests that all the groups experienced sensorimotor adaptation and continuously updated their internal models to match the intended pitch. The aftereffect systematically occurred in the downward shift MA1, suggesting that the online recalibration is task-dependent and related to the closeness of intended pitch and shifted pitch.

Consider the magnitude of compensation in the FAF trials. L2 beginners demonstrated significantly larger compensation (70 cents on average) than L2 advanced learners (61 cents on average) and native Mandarin speakers (56 cents on average) did, while the latter two were not significantly different. This group difference suggests that speakers with well-established internal models for pitch due to years of tonal language experience (i.e., native Mandarin speakers and L2 advanced learners) may be able to suppress pitch-shift responses. This is consistent with the findings in previous literature where singers and Mandarin speakers showed suppressed pitch-shift responses compared to nonsingers and non-native tonal speakers under the constant perturbation (Jones and Keough 2008; Ning et al. 2014; Ning et al. 2015). The suppression indicates that speakers may rely more on feedforward control than on feedback when the speech motor commands have been stored through years of practice (Guenther 2006). The fact that feedforward control outweighs feedback explains why native speakers and L2 advanced learners were less susceptible to pitch errors (i.e., perturbation in auditory feedback) than L2 beginners. Our L2 advanced learners had 5 years of acquisition on average, while L2 beginners only had 2 years of acquisition. This may imply that an entrenched internal model for pitch could require up to 5 years of language learning.

However, the pitch-shift responses cannot be fully suppressed. In other words, the compensation could not be exactly 0 cent. Generally, the magnitude of compensation peaks at 50 cents no matter how large the perturbation is (Larson 1998). The learning paradigm with a sequence of baseline, pitch-shift perturbation, and back-to-normal auditory feedback seems to result in larger pitch-shift response magnitudes. In Jones and Keough (2008), the F0 values were shifted down 100 cents in the FAF trials. The singers' pitch-shift responses in the FAF trials were 60-80 cents, whereas nonsingers' responses were 80-100 cents (almost perfect compensation). In the current study, close to perfect compensation occurred in the block 52-61 of the downward shift MA1 (94 cents) and L2 beginners' AH (95 cents). Keough et al. (2013) argue that shifting pitch from vocal onset would make compensatory responses more difficult to suppress than shifting pitch in the middle of an utterance. This research presenting pitch perturbation from vocal onset also supports Keough et al.'s (2013) argument on the degree of compensation. One anonymous reviewer argued that when there is no time lag between the onset of utterance and the onset of perturbed auditory feedback, auditory feedback may outweigh feedforward control. This is a valid concern and also evident in this study. Heavy reliance on auditory feedback will make speakers generate larger pitch-shift responses (in our case, close to perfect compensation when shifting pitch from vocal onset) to compensate for the mismatch between intended pitch and heard pitch.

Sensorimotor adaptation appeared in the FAF trials for all the groups. Participants can continuously update the internal models for pitch by remapping the auditory-motor representations in the production of AH and MA1. The online recalibration made their F0 values close to the F0 values produced in the previous trial, as shown in the gradual increment or decrement of pitch-shift responses in the FAF trials. This study serves as an initial attempt to examine L2 learners' sensorimotor adaptation. Since all speakers demonstrated sensorimotor adaptation, it suggests that pitch-shift responses are automatic and less likely to be subject to volitional control (Keough et al. 2013).

Although all speakers showed compensation and sensorimotor adaptation, there was a significant interaction between GROUP and TASK. L2 advanced learners and native Mandarin speakers displayed larger compensation in MA1 than AH, while L2 beginners compensated more in AH than in MA1. Chen et al. (2007) argue that pitch-shift responses in constant perturbation were task-dependent, where response magnitudes were larger in the speech condition (32 cents) than in the sustained vowel condition (22 cents). In this study, only L2 advanced learners and native Mandarin speakers demonstrated the same pattern (MA1 > AH) as argued in Chen et al. (2007). It is likely that L2 beginners did not have entrenched internal models for lexical tones and thus the pitch-shift responses in MA1 (linguistically meaningful speech) were not larger than those in AH (nonlinguistic vocalization).

The interaction between GROUP and DIRECTION on the compensation and sensorimotor adaptation is comparatively puzzling. It is unclear why L2 advanced learners did not show larger compensation in the downward shift condition than in the upward shift condition, which seems to be a trend across speakers and conditions (see Figure 2). Previous literature has demonstrated that downward shifts in general elicited larger compensatory responses than upward shifts (Chen et al. 2007; Liu et al. 2011). Why the DIRECTION effect is missing in our L2 advanced learners requires further research.

Consider the aftereffect by comparing the baseline trials and the first ten test trials. All three groups demonstrated the significant aftereffect in the downward shift MA1, but not in the upward shift MA1. The Mandarin syllable MA1 has a high level pitch. The downward shift made the tone considerably deviate from the intended high pitch. Therefore, greater effort to correct the mismatch was made not only in the FAF trials (i.e., larger compensatory responses) but also in the test trials (i.e., significant aftereffect). In comparison, the upward shift was manifested in the same direction as the intended pitch. The closeness of intended pitch and upward shifted pitch required less effort in correction. Thus, in the upward shift MA1, participants' voice F0 values soon returned to the baseline levels when auditory perturbation was removed. This suggests that the online recalibration of the relationships between auditory input and motor command is not only concerned about what the vocalization task is (either speaking or vowel prolongation) but also concerned about the degree of

mismatch that has been generated by comparing the expected voice output and actual voice output. Up until now, it is not clear whether the aftereffect would continue if repeated exposure to pitch perturbation (i.e., practice) would permanently change the sensorimotor representations. If so, we may be able to assist second language learners in reshaping the internal models for pitch. This requires further research.

Another possible explanation for the aforementioned discrepant aftereffect (proposed by one anonymous reviewer) is that the participants (particularly for males) may initiate at a lower pitch of their pitch range in the baseline of MA1, and thus in face of upward shifts, there was no much space for the compensatory responses to go further down in pitch. When pitch perturbation was removed in the test phase, no significant difference between the baseline and the test phase (i.e., no significant aftereffect) would be observed. Since we did not test the participant's pitch range, whether this hypothesis holds requires further research.

This study found no aftereffect in AH (except for L2 advanced learners' upward shift AH), which is inconsistent with the previous results (Jones and Keough 2008; Keough and Jones 2009; Keough et al. 2013). One major difference between the current research and the previous literature is that we did not ask participants to vocalize at a particular musical note, while previous research all asked both singers and nonsingers to vocalize at an absolute pitch value (such as 392 Hz (G4) or 349 Hz (F4) in Jones and Keough (2008); 349 Hz (F4), 392 Hz (G4), or 440 Hz (A4) in Keough and Jones (2009); 220.00 Hz (A3), 246.94 Hz (B3), 293.66 Hz (D4) or 329.63 Hz (E4) in Keough et al. (2013)). Participants in the current study used their comfortable steady pitch to vocalize /a/. They may randomly select a pitch value in the vocalizations. The random selection interfered with the remapping of auditory-motor representations. Hence, the absence of aftereffect in AH may again support the fact that the online recalibration is task-dependent, depending on whether vocal accuracy is required. It's unclear why L2 advanced learners showed aftereffect (only) in the upward shift AH. It is likely they tried to vocalize at a steady and constant low pitch in AH.

In sum, the suppressed responses in the FAF trials were found in L2 advanced learners and native Mandarin speakers, but not in L2 beginners. However, the aftereffect in the test trials was discovered in all the groups. These two findings can be

explained by the internal models. Speakers with entrenched internal models for pitch (as evident in L2 advanced learners and native Mandarin speakers) take the advantage of feedforward commands to monitor the vocal output and thus are less susceptible to the error generated in the feedback loop. Our L2 beginners, on the other hand, are more susceptible to the auditory perturbation probably because they rely more on the auditory feedback signals to monitor the voice F0. Internal models can be recalibrated online in all the groups; thus, the update function is automatic, not subject to language experience. The aftereffect in the downward shift MA1 for all the groups suggests that L2 learners, even for the beginners, have a clear pitch target for Mandarin high-level tone. For future study, it would be interesting to investigate whether L2 learners have clear pitch targets for other Mandarin tones, which could be more challenging to learn.

5. Conclusions

Results from the present study show that L2 advanced learners and native Mandarin speakers had suppressed compensatory responses compared to L2 beginners. This suggests that the former two groups had reduced reliance on auditory feedback during speech production. Remapping the auditory-motor representations was observed online in all speakers, suggesting that compensation and sensorimotor adaptation were automatic. The absence of aftereffect in AH and in the upward shift MA1 indicates that the online recalibration is task-dependent (present when vocal accuracy is required) and related to the closeness of intended pitch and shifted pitch (present when there is a deviation).

References

- Burnett, Theresa A., Marcia B. Freeland, and Charles R. Larson. 1998. Voice F0 responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America* 103.6: 3153-3161.
- Burnett, Theresa A., and Charles R. Larson. 2002. Early pitch-shift response is active in both steady and dynamic voice pitch control. *Journal of the Acoustical Society of America* 112.3: 1058-1063.
- Callan, Daniel E., Jeffery A. Jones, Akiko M. Callan, and Reiko Akahane-Yamada. 2004. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage* 22: 1182-1194.
- Camacho, Arturo, and John G. Harris. 2008. A sawtooth waveform inspired pitch estimator for speech and music. *Journal of Acoustic Society of America* 124.3: 1638-1652.
- Chen, Stephanie H., Hanjun Liu, Yi Xu, and Charles R. Larson. 2007. Voice F0 responses to pitch-shifted voice feedback during English speech. *Journal of the Acoustical Society of America* 121.2: 1157-1163.
- Guenther, Frank H. 1995. Speech sound acquisition coarticulation and rate effects in a neural network model of speech production. *Psychological Review* 102: 594-621.
- Guenther, Frank H. 2006. Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders* 39.5: 350-365.
- Hain, Timothy C., Theresa A. Burnett, Swathi Kiran, Charles R. Larson, Shajila Singh, and Mary K. Kenney. 2000. Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research* 130: 133-141.
- Hain, Timothy C., Theresa A. Burnett, Charles R. Larson, and Swathi Kiran. 2001. Effects of delayed auditory feedback (DAF) on the pitch-shift reflex. *Journal of the Acoustical Society of America* 109.5: 2146-2152.
- Jones, Jeffery A., and Dwayne Keough. 2008. Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research* 190: 279-287.
- Jones, Jeffery A., and Kevin G. Munhall. 2005. Remapping auditory-motor representations in voice production. *Current Biology* 15: 1768-1772.

- Jordan, Michael I., and David E. Rumelhart. 1992. Forward models Supervised learning with a distal teacher. *Cognitive Science* 16: 307-354.
- Kawato, Mitsuho. 1999. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology* 9.6: 718-727.
- Keough, Dwayne, Colin Hawco, and Jeffery A. Jones. 2013. Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neuroscience* 14: 25.
- Keough, Dwayne, and Jeffery A. Jones. 2009. The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *Journal of Acoustic Society of America* 126.2: 837-846.
- Lalazar, Hagai, and Eilon Vaadia. 2008. Neural basis of sensorimotor learning Modifying internal models. *Current Opinion in Neurobiology* 18: 573-581.
- Larson, Charles R. 1998. Cross-modality influences in speech motor control The use of pitch shifting for the study of F0 control. *Journal of Communication Disorders* 31: 489-503.
- Larson, Charles R., Theresa A. Burnett, Jay J. Bauer, Swathi Kiran, and Timothy C. Hain. 2001. Comparison of voice F0 responses to pitch-shift onset and offset conditions. *Journal of the Acoustical Society of America* 110.6: 2845-2848.
- Larson, Charles R., Theresa A. Burnett, Swathi Kiran, and Timothy C. Hain. 2000. Effects of pitch-shift velocity on voice F0 responses. *Journal of the Acoustical Society of America* 107.1: 559-564.
- Liu, Hanjun, Michelle Meshman, Roozbeh Behroozmand, and Charles R. Larson. 2011. Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback. *Clinical Neurophysiology* 122: 951-957.
- Liu, Hanjun, Yi Xu, and Charles R. Larson. 2009. Attenuation of vocal responses to pitch perturbations during Mandarin speech. *Journal of the Acoustical Society of America* 125.4: 2299-2306.
- Max, Ludo, Frank H. Guenther, Vincent L. Gracco, Satrajit S. Ghosh, and Marie E. Wallace. 2004. Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders* 31: 105-122.
- Ning, Li-Hsin, Torrey M. Loucks, and Chilin Shih. 2015. The effects of language

- learning and vocal training on sensorimotor control of lexical tone. *Journal of Phonetics* 51: 50-69.
- Ning, Li-Hsin, Chilin Shih, and Torrey M. Loucks. 2014. Mandarin tone learning in L2 adults: A test of perceptual and sensorimotor contributions. *Speech Communication* 63-64: 55-69.
- Perkell, Joseph, Melanie Matthies, Harlan Lane, Frank Guenther, Reiner Wilhelms-Tricarico, Jane Wozniak, and Peter Guiod. 1997. Speechh motor control Acoustic goals saturation effects auditory feedback and internal models. *Speech Communication* 22: 227-250.
- Scheerer, Nichole E., and Jeffery A. Jones. 2012. The relationship between vocal accuracy and variability to the level of compensation to altered auditory feedback. *Neuroscience Letters* 529.2: 128-132.
- Wolpert, Daniel M., Zoubin Ghahramani, and Michael I. Jordan. 1995. An internal model for sensorimotor integration. *Science* 269.5232: 1880-1882.
- Xu, Yi, Charles R. Larson, Jay J. Bauer, and Timothy C. Hain. 2004. Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. *Journal of the Acoustical Society of America* 116.2: 1168-1178.

[審查：2018.3.16 修改：2018.6.5 接受：2018.6.30]

甯俐馨

Li-Hsin NING

10610 臺北市和平東路一段 162 號 國立臺灣師範大學英語系

Department of English

National Taiwan Normal University

No.162, Sec. 1, Heping E. Rd., Taipei City 10610, Taiwan

lihsin@ntnu.edu.tw

以華語為第二語言學習者於聽覺反饋干擾中 產生的發聲動作適性化

甯俐馨

國立臺灣師範大學英語學系

摘要

本研究旨在探討以華語為第二語言學習者的發聲動作適性化。本研究使用帶有適性化音頻變調的典範（音頻逐漸上升或下降）來檢視初學者、進階班學生、華語母語人士於母音 /a/ 與「媽」發聲時，對於聽覺反饋干擾的反應。研究結果顯示相較於初學者，進階班學生與華語母語人士較少仰賴聽覺反饋的資訊來監控音頻高低。這三組受試者皆發現感覺動作的適性化，表示內在模組的即時校正功能是自動化的。然而，內在模組的即時校正功能卻是任務導向的，出現與否取決於發聲過程當中是否要求音頻準確度、以及音頻變調的方向是否趨近於預期的音頻高度。

關鍵詞：感覺動作適性化 音頻變調典範 內在模組 以華語為第二語言學習者