

2010

The Role of Auditory Feedback on the Control of Voice Fundamental Frequency (F_0) while Singing

Dwayne Nicholas Keough
Wilfrid Laurier University

Follow this and additional works at: <http://scholars.wlu.ca/etd>

 Part of the [Cognitive Psychology Commons](#)

Recommended Citation

Keough, Dwayne Nicholas, "The Role of Auditory Feedback on the Control of Voice Fundamental Frequency (F_0) while Singing" (2010). *Theses and Dissertations (Comprehensive)*. 1090.
<http://scholars.wlu.ca/etd/1090>

This Dissertation is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact scholarscommons@wlu.ca.

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]





Library and Archives
Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-64397-6
Our file *Notre référence*
ISBN: 978-0-494-64397-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

■◆■
Canada

THE ROLE OF AUDITORY FEEDBACK ON THE CONTROL OF VOICE
FUNDAMENTAL FREQUENCY (F_0) WHILE SINGING

by

Dwayne Nicholas Keough

Bachelor of Arts (Honours), Memorial University of Newfoundland, 2003

Master of Science, Wilfrid Laurier University, 2006

DISSERTATION

Submitted to the Faculty of Science

In partial fulfillment of the requirements for

Doctor of Philosophy in Psychology

Wilfrid Laurier University

2010

Dwayne N. Keough © 2010

Abstract

Whether we are learning how to play a new instrument, song, or even learn a second language, the nervous system relies on various forms of sensory feedback to establish task-specific sensorimotor representations. Over time, the plasticity of the nervous system permits neural reorganization and the formation of an 'internal model'. It has been suggested that internal models represent neural maps of skilled movement that store the relationship between the motor commands, environment and sensory feedback responsible for their production. These internal representations are often investigated by altering a particular aspect of the sensory feedback associated with a given task. Arguably the most influential contribution to F_0 control during speech production is from auditory feedback. For instance, using the frequency-altered feedback (FAF) paradigm (where the pitch of participants' voices is shifted in frequency) results show that participants will compensate by adjusting their vocal productions in the opposite direction of the perturbation. Moreover, aftereffects are often observed (subsequent responses err in the direction of compensation) when auditory feedback is unexpectedly returned to normal following a series of fixed pitch-shift manipulations. Thus, the evidence from studies on speech production suggests that an internal model represents voice fundamental frequency (F_0). However, there is little evidence to suggest that this is also true for vocal control while singing. As a result, the purpose of this dissertation is to examine whether acoustic-motor representations (internal models) control voice F_0 .

Chapter 1 examined trained singers and untrained singers (nonsingers) sensitivity to subtle manipulations in auditory feedback. Participants produced the consonant-vowel /ta/ while receiving auditory feedback that was shifted up and down in

frequency. Results showed that singers and nonsingers compensated to a similar degree when presented with frequency-altered feedback (FAF), however, singers' F_0 values were consistently closer to the intended pitch target. Moreover, singers initiated their compensatory responses when auditory feedback was shifted up or down 6 cents or more, compared to nonsingers who began compensating when feedback was shifted up 26 cents and down 22 cents. Additionally, examination of the first 50 ms of vocalization indicated that participants commenced subsequent vocal utterances, during FAF, near the F_0 value on previous shift trials. Interestingly, nonsingers commenced F_0 productions below the pitch target and increased their F_0 until they matched the note. We conclude that singers and nonsingers rely on an internal model to regulate voice F_0 , but singers' models appear to be more sensitive in response to subtle discrepancies in auditory feedback.

Evidence from arm reaching studies suggests that the motor system can acquire multiple internal models, which has been argued to allow an individual to adapt to different perturbations in diverse contexts. In Chapter 2 we show that trained singers can rapidly acquire multiple internal models that accommodate different perturbations to ongoing auditory feedback. Participants heard three musical notes and reproduced them in succession. Adaptation was observed immediately following vocal onset when participants were gradually exposed to altered feedback. Aftereffects were target specific and did not influence vocal productions on subsequent trials. Interestingly, the target notes in Chapter 2 could have served as a contextual cue. Indeed, when the target notes could no longer serve as a contextual cue we observed evidence for trial-by-trial adaptation. These findings indicate that the brain is exceptionally sensitive to deviations between auditory feedback and the predicted consequence of a motor

command during vocalization. Moreover, these results indicate that when contextually cued the vocal control system may maintain multiple internal models that are capable of independent modification during different tasks or environments.

Auditory feedback has been shown to be influential in the maintenance and control of voice F_0 . The purpose of Chapter 3 was to address whether task instructions could influence the compensatory responding and sensorimotor adaptation that has been previously found when participants are presented with a series of FAF trials. Trained singers and nonsingers were informed that their auditory feedback would be manipulated in pitch while they sang the target vowel [ah/] (as in 'pop'). Participants were instructed to either 'compensate' or 'ignore' the changes in auditory feedback. Auditory feedback manipulations persisted for the entire vocal production and were either gradually presented ('ramp') in -2 cent increments down to -100 cents (1 semitone) or were suddenly ('constant') shifted down by 1 semitone. Results indicated that singers and nonsingers could not suppress the compensatory responses to FAF, nor could they reduce the sensorimotor adaptation observed on both ramp and constant FAF trials. Compared to previous research these data suggest that musical training is effective in suppressing compensatory responses only when FAF occurs between 500-2500 ms following vocal onset. Moreover, our data suggest that compensation and adaptation is automatic and is influenced little by conscious control.

Indeed, regardless of whether we manipulate auditory feedback in small or large increments, for a single music note or multiple notes, or whether we modify the task instructions (ignore or compensate), we have shown and that the result is a change in the sensorimotor representations (internal models) underlying voice F_0 control while singing. One goal has been to demonstrate the plasticity of the nervous system by

examining how changes in singers and nonsingers auditory feedback can alter target specific internal representations. Indeed, producing multiple target notes while singing was shown to require participants to employ unique motor commands for each target. Although the human voice has the potential to be initiated at unpredictable frequencies during vocal onset, we found that trained singers consistently produced vocal pitch at frequencies near the desired target, even in the presence of FAF. Overall, it appears that compensation and adaptation to FAF are automatic, influenced little by 'conscious' control, and are uniquely associated with the motor commands for specific musical targets.

Acknowledgments

Accomplishing this goal, in many ways, has been my most challenging endeavour to date. Firstly, I would like to thank my supervisor, Dr. Jeffery Jones, for his support and guidance over the years. He has offered me invaluable advice and I am fortunate to consider him a friend and mentor. I would also like to thank my previous supervisors, Dr. Angelo Santi and Dr. David Brodbeck (masters and undergraduate, respectively) for their direction and for showing me that I had the ability to be where I am today. I would also like to thank all three of my supervisors for kicking my butt when I needed it. I am truly grateful for their belief in me, even when I didn't believe in myself.

I would also like to acknowledge my fellow graduate students whom I have had the pleasure to work with and get to know over the years. I consider myself fortunate to have worked with you and appreciate your discussions on my own research as well as others. My fellow lab alumni (especially Colin, Michelle, & Farina), you made coming to the lab exciting everyday and I wish you all the best of luck with your future research endeavours and careers. To my friends, thank you for your support and conversation. It was nice to hear what was going on in the 'real world'. I own a special thanks to Jason, Gisele, Peggy, Ron, Aaron, Jon, Amy and Jeff.

This section would not be complete if I forgot the most important people in my life, my family. To my parents, your support and guidance over the years was greatly appreciated. Even though it was difficult to ask for help at times you always answered the call. I want you to know that I always looked up to you and continue to do so. To my outlaws, you have treated me like your own child even though you don't have to. I want you to know that I would be very proud to call you mom and dad (even though you prefer Gerry and Michelle). You are remarkable parents, grandparents and I am glad to

have you in our lives. To my brothers and sister-by-law thank you for always being there. No matter how challenging life was you were always there to lend a hand or just listen to my rants. I will promise to do the same for you when you need me, no matter where or when. I would also like to thank the entire Keough, Tobin, Collins, and Collins crew for their love and support. I am fortunate to have my grandmother with me and I would like her to know that I cannot express my gratitude for her belief in me.

Grandparents are special people, and I wish they were all here today to celebrate my accomplishment.

Lastly, I dedicate this dissertation to my beautiful wife, Stephanie, and our son Thomas (and the baby-to-be). I wake up every day thrilled that I have my best friend by my side and our wonderful little man near by. You are the first thing I think about every day and the last before I sleep. You are the reason I am here and I am positive that I would not have made it to the end without your love, support and patience over the years. My heart is yours, as it always has been.

Table of Contents

Abstract	ii
Acknowledgments	vi
Table of Contents	viii
List of Figures	ix
General Introduction.....	1
Chapter 1	29
Background	31
Methods	38
Results	41
Discussion	51
Chapter 2	62
Background	65
Methods	67
Results	72
Discussion	76
Chapter 3	90
Background	92
Methods	100
Results	106
Discussion	112
General Discussion.....	128
References	152

List of Figures

Figure 1.1:	60
Figure 1.2:	61
Figure 2.1:	87
Figure 2.2:	88
Figure 2.3:	89
Figure 3.1:	125
Figure 3.2:	126
Figure 3.3:	127

General Introduction

An Overview

In general, I am interested in motor control, particularly how some individuals seem to possess a greater ability to control their behaviour than others. For instance, why are some individuals ('singers') capable of 'carrying a tune' whereas others seem to be 'tone-deaf'? Is singing proficiency attributable to the fact that individuals were born with this ability, or was this something that was acquired through years of training? Moreover, what is the influence of sensory feedback (e.g., auditory, kinesthetic, vision) on motor control? How important is auditory feedback on singers and nonsingers ability to control the pitch of their voice while singing? Indeed, these are just a few of the questions that have driven my passion for research on the motor control of singing. The focus of this dissertation is to examine how sensory feedback is used by the nervous system in singers and nonsingers to establish and maintain internal representations ('internal models') for motor control.

Regardless of whether one is learning to speak a new language, sing a new song, or learning a new sport, various forms of sensory feedback assist with the formation of task-specific sensorimotor representations. In time, the plasticity of the nervous system allows for neural reorganization and the formation of an 'internal model'. The majority of research supporting the internal model hypothesis originates from studies manipulating the sensory feedback associated with arm-reaching movements. Indeed, the prevailing hypothesis for the control of limb dynamics (Wolpert & Kawato, 1998), and the control of speech (Houde &

Jordan, 1998; Guenther & Perkell, 2004; Jones & Munhall, 2000, 2005) and singing (Jones & Keough, 2008) is that internal models regulate motor movements. Essentially, researchers are interested in observing the motor plans for newly acquired or well-established movements following various forms of feedback manipulations. Results have shown that participants compensate for changes in feedback by altering their movements in the opposite direction of the perturbation (e.g., Burnett, Freedland, Larson, & Hain, 1998; Burnett, Senner, & Larson, 1997; Burnett & Larson, 2002; Donath, Natke, & Kalveram, 2002; Elman, 1981; Ghahramani & Wolpert 1997; Jones & Keough, 2008; Jones & Munhall, 2000, 2005; Kalenscher, Kalveram, & Konczak, 2003; Larson, 1998; Larson, Burnett, Kiran, & Hain, 2000; Natke, Donath, & Kalveram, 2003; Sainburg, Ghez, & Kalakanis, 1999; Shadmehr & Moussavi, 2000; Shadmehr & Mussa-Ivaldi, 1994; Toyomura et al., 2007). Moreover, sensorimotor adaptation or aftereffects have been found when feedback suddenly returns to normal following a series of perturbation trials (Ghahramani & Wolpert, 1997; Jones & Keough, 2008; Jones & Munhall, 2000, 2005; Shadmehr & Moussavi, 2000; Shadmehr & Mussa-Ivaldi, 1994). That is, participants initiate movements as if they were expecting altered feedback. As a result, aftereffects are thought to represent the sensorimotor recalibration of internal models for motor control.

The purpose of this dissertation is to further our understanding of the role of auditory feedback on the acoustic-motor representations (internal models) controlling voice fundamental frequency (F_0). Indeed, increasing or decreasing the amplitude (loudness) of vocal productions (Bauer, Mittal, Larson, & Hain,

2006), changing formant frequencies (Houde & Jordan, 1998; Munhall, MacDonald, Byrne, & Johnsrude, 2009; Purcell & Munhall, 2006), delaying auditory feedback (Yates, 1963), and altering participants fundamental frequency (pitch) (e.g., Burnett et al., 1997, 1998; Burnett & Larson, 2002; Elman, 1981; Larson, Burnett, Bauer, Kiran, & Hain, 2001; Hain et al., 2000; Jones & Keough, 2008; Jones & Munhall, 2000, 2002, 2005; Zarate & Zatorre, 2005, 2008) are some of the methods used to address this question. However, the primary focus of this dissertation will be how frequency-altered feedback (FAF) affects voice F_0 control while singing. For example, do singers and nonsingers compensate for frequency-altered feedback, and if so, to an equal degree? Are similar levels of sensorimotor adaptation found following FAF trials? Do aftereffects generalize to other notes that were not subjected to FAF, or are there multiple internal models that correspond to specific musical targets? Can singers and nonsingers ignore FAF when instructed to do so? These are the questions that my dissertation research has addressed. Chapter 1 examined singers and nonsingers vocal responses to sudden and persistent changes (-100 cents; 1 semitone) in auditory feedback. Chapter 2 examined singers and nonsingers sensitivity to subtle changes in auditory feedback manipulations (2 cent changes across trials to 100 cents). Chapter 3 examined whether multiple internal models exist for vocal control and chapter 4 will discuss singers and nonsingers ability to follow instructions to compensate and ignore FAF manipulations.

Internal Models for Motor Control: A Brief Review

Whether individuals are performing movements in novel contexts or are executing well rehearsed behaviours, the sensorimotor feedback associated with these tasks are thought to form or derive from an 'internal model' responsible for motor control (e.g., Wolpert & Kawato, 1998). Internal models are believed to represent neural maps of skilled movement that store the relationship between the motor commands, environment and sensory feedback responsible for their production (Desmurget & Grafton, 2000; Flanagan & Wing, 1993; Shadmehr & Mussa-Ivaldi, 1994). Moreover, internal models are thought to represent the plasticity of the nervous system, which is believed to be the neural basis for adaptive behaviour (Kawato, Furukawa, & Suzuki, 1987). Adaptive behaviour is a basic function that is central to the successful performance of the most basic tasks necessary for survival (Gidley Larson, Bastian, Donchin, Shadmehr, & Mostofsky, 2008). For instance, it is particularly advantageous that individuals can alter their motor plans (e.g., walking while holding an object) to overcome various internal (e.g., fatigue) and external (e.g., holding a child) environmental demands. Indeed, Gidley Larson et al. (2008) argue that adaptation is so crucial for survival that adaptation in individuals with autism spectrum disorder appears to be preserved at the expense of other cerebellar functions. Although considerable evidence supports the notion that the cerebellar cortex can acquire internal models through motor learning (e.g., Gomi, et al., 1998; Imamizu et al., 2000; Kawato & Gomi, 1992; Kitazawa, Kimura, & Yin, 1998; for a review see Ito,

2008), the intention of this section is to introduce the basic premise of what we mean by an 'internal model' and how it applies to motor control.

An internal model encompasses the entire sequence of neural events that represent a motor movement; from premotor sequence selection, to the online monitoring of current productions, and ultimately the formation of a neural representation for the completed movement. Kawato, Furukawa and Suzuki (1987) provide a very specific account to describe how this is computed: (1) The cortex sends the desired motor plan to the motor cortex. (2) The motor command is then computed, which estimates the torque to be generated by the muscles. (3) The motor command is then sent to the muscles via the spinal motoneurons. (4) The musculoskeletal system interacts with the environment that realizes some type of motor pattern. (5) The actual plan (current state of the effector) and time derivative is measured by proprioceptors, which is then sent back to the motor cortex via a negative feedback loop. (6) Feedback control is now able to be computed using error. (7) Spinocerebellum magnocellular part of the red nucleus system receives information about the movement as afferent input from proprioceptors as well as from efference copy. (8) Within the cerebellum, an internal model is formed. (9) Once acquired through motor learning, the model can provide an estimate of the actual movement.

In more simple terms, a motor command is computed ('forward model') based on the information available (e.g., vision, state of limb) prior to initiating a movement (e.g., grasping a cup). Forward models are believed to predict upcoming states of the system based on the current state and active motor

command. Thus, the initial phase of the movement (e.g., between 0-100 ms) occurs via open-loop control, or without relying on sensorimotor feedback. Essentially, the brain selects and implements a motor plan given the information available. During mid to late production, if errors are detected then the brain is able to initiate compensatory responses (e.g., between 150-250 ms), which occur via closed-loop control, such that participants can use feedback to alter their motor plans (e.g., within the same or across successive productions) and accomplish the task (e.g., pick up the cup). Over time and with repeated exposure, the brain establishes an optimal strategy, or internal model, for that particular movement. This optimal strategy refers to how the brain learns to produce motor commands that minimize costs and maximize rewards (Krakauer & Shadmehr, 2007). Overall, internal models are proposed to explain how feedforward controllers may account for the numerous motor commands produced during diverse motor contexts experienced by individuals on a daily basis (Wolpert and Kawato, 1998).

The theory of internal models proposes that two types of controllers exist (forward and inverse models). Note that the inverse model is not a model of the external world, rather it is a model of information processing carried out in other brain regions (Kawato et al., 1987). Essentially, Kawato et al. describe an inverse model as the dynamic system whose input and output are inverted (e.g., trajectory is input and motor command is output). Typically, optimal performance requires both forward and inverse models. For example, during reaching tasks participants generally hold a robotic arm and attempt to move a cursor on a

computer screen to a target location, within a fixed period of time. If done correctly, the target will explode otherwise if the movement is too slow the participants are given a warning to make the movement more quickly. When the force is applied to the arm during the reaching movement, the best strategy is not a nearly straight trajectory to reach a target. The optimal cursor trajectory is slightly curved. That is, overcompensate early in the movement and under compensate near the end. The reason this is optimal, according to Krakauer and Shadmehr (2007), is that feedback allows one to correct errors occurring early in the movement but not errors occurring late. Therefore, it is easier to make the movement with the force, than to oppose it. It was also noted that the inverse model would not be formed/employed if catch trials were present (Krakauer & Shadmehr, 2007). Thus, it seems that the overcompensation strategy is only optimal in predictive instances, where one is certain that the force will be experienced. If the task is unpredictable, as is the case in a number of FAF studies examining the importance of auditory feedback on vocal motor control (e.g., Burnett et al., 1997, 1998; Donath et al., 2002; Hain et al., 2000; Larson, 1998; Larson, Burnett, Kiran, & Hain, 2000; Liu and Larson, 2007), then the forward model will be employed. Given the presumption that feedback manipulations act on the forward model, and not the inverse model, the remainder of this dissertation will only focus on forward models of motor control. Therefore, forward model will be appear as 'internal model' throughout the remainder of the document.

Sensory Consequences of Action: The Efference Copy

An important component of the theory of internal models is the sensory consequences of actions. It is hypothesized that forward models include an 'efference' comparator, or a copy of the motor plan, which is sent to the sensory cortex to predict the sensory consequences of our actions (Nowak, Topka, Timmann, Boecker, & Hermsdörfer, 2007). The critical function of the efference copy of our motor plan is to assist with the differentiation between self-generated and externally-generated stimuli (Blakemore, Wolpert, & Frith, 2000). For instance, if the predicted outcome of a motor plan is congruent with the actual feedback received from the movement, no error is registered. Conversely, if the sensory feedback does not match the consequences of the predicted motor plan, then an error is detected, which results in the brain eliciting compensatory responses to reduce the error on subsequent trials. In the case of vocalization, the increased activation found when auditory feedback has been manipulated, relative to when it is unaltered, is thought to represent a violation in the expected sensory consequences of the motor plan and the actual results. Thus, the forward model's efference copy is important in not only helping individuals differentiate between self-generated movements from externally-generated movements, but it is also thought to be related to the compensatory responses that occur to manipulated feedback in humans.

Evidence for Cortical Suppression During Vocalization

Studies using event-related potentials (ERP) and magnetoencephalography (MEG) have found suppression in cortical activity occurring approximately 100 ms (N100 & M100, respectively) following vocalization onset in humans (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Heinks-Maldonado, Nagarajan, & Houde, 2006) and in marmoset monkeys (Eliades & Wang, 2008). For instance, Heinks-Maldonado and colleagues (2005, 2006) investigated whether a forward model operates in the auditory system that causes a dampened sensory experience when you hear your own unaltered voice, your own pitch-shifted voice (200 cents, 2 semitones), or an alien unaltered or pitch-shifted voice. During speech production, it was found that the M100 (an early sensory detection component generated in the auditory cortex) was maximally suppressed to participant's own unaltered voice relative to the pitch-shifted and alien feedback conditions (Heinks-Maldonado et al., 2006). On the other hand, when participants were required to simply listen to speech, the various forms of feedback (own verses alien) did not yield M100 amplitude differences compared to when participants actively vocalized during the task. Moreover, similar results were obtained using ERPs, except the component of interest was the N100 (Heinks-Maldonado et al., 2005). It appears that the cortical suppression observed in the auditory cortex is potentially a consequence of a precise forward model mechanism. The following section will address the role of sensory feedback on motor control and how compensatory responses have demonstrated the recalibration of an internal model.

The Role of Sensory Feedback for Motor Control: Compensatory Responding

There is little doubt that sensory feedback is essential for the acquisition and maintenance of precise motor control (e.g., Burnett et al., 1997, 1998; Elman, 1981; Ghahramani & Wolpert, 1997; Houde & Jordan, 1998; Kalenscher et al., 2003; Jones & Keough, 2008; Jones & Munhall, 2000, 2002, 2005; Larson et al., 2008; Munhall et al., 2009; Sainburg et al., 1999; Shadmehr & Moussavi, 2000; Shadmehr & Mussa-Ivaldi, 1994). Indeed, Finney and Palmer (2003) demonstrated that auditory feedback is important during the acquisition of a novel musical piece in trained pianists. Moreover, receiving auditory feedback has been shown to be crucial in learning to produce speech (Oller & Eilers, 1988), and for the maintenance of precise articulation across the lifespan (Cowie & Douglas-Cowie, 1992). In regards to vocal productions, previous research has found that delaying auditory feedback (Yates, 1963) disrupts the quality and fluency of vocal productions (e.g., stuttering). However, changing auditory feedback typically results in compensatory responses (in the opposite direction of the manipulation) in speakers' ongoing vocal productions. For instance, have you ever noticed that when you are trying to talk to someone in a crowded environment you suddenly and automatically raise your voice in an attempt to facilitate communication? In 1911, Etienne Lombard predicted that when speakers are exposed to loud noise they spontaneously increase the intensity of their voice, which is essentially stating that they were compensating for the noise (Lane & Tranel, 1971). In honour of this important finding, the effect was appropriated referred to as the 'Lombard' effect.

In regards to compensation and adaptation, an interesting observation was found in Held's (1965) classic study following the prolonged exposure to prism goggles. Held had participants wear prism goggles that horizontally displaced their visual field. Initially, errors were made in the direction that the prism goggles were altering their vision. Over time, participants compensated by making movements in the opposite direction of the displacement, such that accuracy and the duration required to complete the task was similar to baseline performance (without the goggles). When the goggles were removed, participants made movements that were in the direction of compensation (aftereffects). This suggests that participants were anticipating the visual displacement on trials following the removal of the prism goggles. Interestingly, Held failed to find adaptation when participants passively viewed a scene while wearing prism goggles (they were pushed around by the experimenter while standing on a platform). Thus, due to the fact that even though the eye received similar information during the active and passive viewing conditions, it appears that the crucial connection between sensory input and motor output was missing in the passive condition (Held, 1965). Thus, movement alone without the opportunity for error recognition does not suffice to produce adaptation, rather Held argues that the movement must be self-produced.

Other examples demonstrating compensatory responses to auditory feedback manipulations include presenting participants with increasing masking noise or decreasing side-tone amplitude (Bauer et al., 2006; Lane & Tranel, 1971), shifting formant frequencies (Houde & Jordan, 1998; Munhall et al., 2009;

Purcell & Munhall, 2006), or shifting participants' fundamental frequency (F_0) (Burnett et al., 1997, 1998; Elman, 1981; Jones & Munhall, 2000, 2002, 2005; Jones & Keough, 2008; Zarate & Zatorre, 2008). Given that this dissertation will focus on the role of auditory feedback on voice F_0 control while singing, the following section will introduce what the FAF paradigm is, how it is used, and the results obtained in previous studies (compensation and sensorimotor adaptation).

Frequency-Altered Feedback: An Overview

Elman (1981) was among the first to investigate the behavioural effects of manipulating the frequency (F_0) of subjects auditory feedback using frequency-altered feedback (FAF). The FAF paradigm typically involves participants producing a nonsense syllable (e.g., /ah/) or a sustained vowel (e.g., /u/) while receiving either normal or altered auditory feedback. In general, the unaltered feedback condition is used to establish a baseline value for which to compare participants' vocal productions while receiving FAF. When participants receive FAF, they are essentially presented with auditory feedback that is higher or lower in pitch than what they are actually producing. In response to this perturbation, the majority of participants will compensate for the discrepancy detected between perception and production by adjusting their F_0 in the opposite direction of the manipulation (e.g., Bauer et al., 2006; Burnett & Larson, 2002; Burnett et al., 1997, 1998; Jones & Munhall, 2000, 2002, 2005; Jones & Keough, 2008; Liu, Zhang, Xu, & Larson, 2007; Natke, Donath, & Kalveram 2003). For instance,

participants will increase their F_0 in response to downward shifts (e.g., one semitone, 100 cents) in auditory feedback, and the same is true for shifts in the opposite direction. However, in rare cases participants will adjust their F_0 in the same direction ('following') as the manipulation (e.g., Burnett et al., 1997, 1998; Burnett & Larson, 2002). On these trials participants will increase their F_0 in response to upward shifts (e.g., 100 cents) in auditory feedback (e.g., pitch), and the same is true of manipulations in the opposite direction. Also, regardless of whether the pitch manipulations coincide with vocal onset (e.g., Heinks-Maldonado et al., 2005; Jones & Keough, 2008, Jones & Munhall, 2000, 2002, 2005) or are presented during mid utterance (e.g., Burnett et al., 1997, 1998; Hawco & Jones, 2009; Larson et al., 2001; Zarate & Zatorre, 2008), participants typically compensate for FAF.

Frequency-Altered Feedback: Compensation using Relative Pitch Targets

Given that numerous FAF studies have been carried out on speech (Burnett et al., 1998; Donath et al., 2002; Jones & Munhall, 2000; Larson, 1998; Larson et al., 2000), participants have only been required to produce a target at a 'relative' frequency. In other words, their vocalizations were not intended to match a specific frequency, as in singing. Interestingly, participants have not exhibited complete levels of compensation during F_0 manipulations during speech production (e.g., Burnett et al., 1997; Donath et al., 2002; Larson, 1998; Larson et al., 2000; Natke et al., 2003). The response magnitudes reported from prior research have ranged from approximately 15-65 cents with an average

response magnitude of around half a semitone, or 50 cents. Indeed, these results have been observed regardless of the direction of the manipulation (upward or downward) and for manipulations varying from 100 to 600 cents. Donath and colleagues (2001) point out that despite the fact that magnitudes of less than one semitone (100 cents) appear to be minute, the levels of compensation fall within natural prosodic variations. For instance, Bosshardt, Sappok, Knipschild, and Holscher (1997) found that voice F_0 can increase about a semitone (100 cents) when participants end a question. Additionally, the mean response latency for compensatory responses has been found to be around 100-150 ms, whereas for following responses it has been found to be around 250-600 ms (Larson, 1998; however see Experiment 2 by Burnett et al., 1998). This implies that the audiovocal system typically detects the discrepancy between perception and production and acts to correct this error (compensate) within approximately 150 ms.

Frequency-Altered Feedback: Compensation using Absolute Pitch Targets

However, unlike speech, singing offers researchers a unique window into the study of F_0 control. Singing involves producing a succession of musical sounds at a particular (absolute) frequency (e.g., 440 Hz, A4), or the number of cycles per second that represents a specific target (a musical note). Thus, singers must maintain accurate vocal control over their F_0 to hit the desired notes. For instance, it is not uncommon for singers to have to match the pitch of their voice to an external reference (that may or may not be producing the same

note) or, conversely, to match the pitch of their voice to a note in the absence of the external referent (when singing a cappella). Sundberg (1987, p. 177) pointed out that trained singers were capable of matching the target (A4, 440 Hz) with an accuracy of less than 1 Hz. Presently there are few studies that have directly examined the role of auditory feedback while singing (Burnett et al., 1997; Jones & Keough, 2008; Natke et al., 2003; Zarate & Zatorre, 2005). Indeed, the results obtained during singing studies using FAF suggest the results are similar to those obtained during speech production: participants compensate for upward or downward manipulations by adjusting their F_0 in the opposite direction of the perturbation.

However, one difference between previous work on speech and singing using FAF was the level of compensation observed. Burnett et al. (1997) reported exemplary data for one trained singer who exhibited near perfect levels of compensation to FAF while a singing musical scale. Moreover, Natke and colleagues (2003) found that participants, who were not trained singers, compensated more during the singing condition (66 cents) than during the speaking condition (47 cents). Additionally, they reported aftereffects, such that the compensatory responses participants made not only lasted longer during the singing task but also persisted into the next trial, such that their post-adaptation F_0 values on initial test trials were higher than their average baseline F_0 values. The fact that singing appears to be more tightly controlled than speaking suggests that it could potentially be the result of the task constraints of singing. For example, notes provide an external reference in which to match voice F_0 and

the note has an absolute pitch value (Burnett et al., 1997; Natke et al., 2003). As a consequence, it appears that using an absolute reference during singing may result in greater levels of compensation to FAF.

Compensation: Unpredictable verses Predictable Manipulations

In regards to the previously mentioned differences in the level of compensatory responding to FAF during speech and singing, it should be noted that some of the studies examining speech were designed to assess compensation (e.g., Burnett et al., 1998; Hain, et al., 2000; Larson et al., 2001), whereas studies interested in singing (Jones & Keough, 2008) were designed to examine sensorimotor adaptation. This is particularly important because in a compensation study the perturbations are often randomly presented, they could occur during mid utterance, the duration of the manipulation may vary (e.g., between 100-500 ms), or multiple perturbations could occur within a single utterance. Conversely, when assessing sensorimotor adaptation the manipulations are typically held constant across several trials, or if the manipulations are progressively increasing or decreasing then the shift values across trials occur in a predictable fashion.

The predictable nature of the task is what has been thought to permit the modification of internal models for motor control. Essentially, the brain anticipates the altered feedback on a particular trial and the motor commands associated with that movement are adjusted to compensate for the error. If the manipulations are not predictable then adaptation is not likely to occur, as the

'best' motor plan is the one that is considered the default (baseline) for the movement. Thus, movements will be initiated from this consistent value and altered when perturbed feedback is randomly encountered.

Random Perturbation of Voice F_0 Results in Two Compensatory Responses

Interestingly, when auditory feedback manipulations are presented randomly following vocal onset during speech production two types of compensatory responses have been reported (Burnett et al., 1997, 1998). For instance, when Burnett et al. (1998) shifted F_0 for intervals between 100 to 500 ms they found that compensatory responses consisted of an early vocal response (VR1) and late vocal response (VR2). The early vocal response was found to have a latency of 100-150 ms post stimulus onset, whereas the late vocal response occurred with a latency of 250-600 ms.

The speed of the vocal response is worth noting because it may identify whether it is 'reflexive' or 'voluntary'. However, using the terms 'reflexive' and 'voluntary' have created a long-standing debate as to whether they are useful scientific concepts, or whether they are prescientific terms that should be discarded. In any event, I will not present the arguments here, but see Prochazka et al. (2000) for a modern take on this debate. Indeed, Burnett and colleagues argue that the first response appears to be relatively automatic. Moreover, they suggest that this response is not strongly influenced by instructions to the participant (e.g., Hain et al., 2000). Hain et al. (2000) examined vocal responses while informing participants to ignore any changes in auditory feedback perceived over the headphones. The pattern of compensatory VR1 responses

observed was similar to those observed when participants were asked to compensate for any perceived changes in feedback. Overall, the first response appears to occur in the opposite direction of the manipulation (compensatory) for small amplitude shifts, and the amplitude of the response is between 25-50 cents, which does not increase for shifts beyond 50 cents in a compensation paradigm (Larson, 1998). As a consequence, Larson's group posits that the VR1 is an automatic feature in the audio-vocal system that is designed to monitor and correct for small errors in voice F_0 feedback. They also posit that the VR2 is under volitional control, as the latency and amplitude of the response differed under various experimental conditions (Hain et al., 2000). Thus, the late response is thought to be a reflection of the voluntary mechanisms employed to modulate voice F_0 while speaking or singing (Larson, 1998).

An important aspect of Larson and colleagues research is that they argue the human voice is rather unresponsive to changes in other types of acoustical signals aside from a person's own voice F_0 (Larson, 1998). If the audio-vocal system did not respond in this way then our vocal productions would be influenced by environmental sounds, resulting in a powerful effect on the control of one's voice F_0 . Moreover, if one's voice F_0 was consistently influenced by other sounds in our environment, then it would stand to reason that we would not possess volitional control of our own voice. Given that this is not the case, the audio-vocal system, according to Larson (1998), appears to be largely designed to correct for small errors in vocal productions, while being capable of largely ignoring other environmental sounds.

Internal Models for Vocal Control: Closed-Loop and Open-Loop Control

It has been proposed by Larson and colleagues (2000) that a closed-loop negative feedback system is operating to compensate for participants anticipated feedback with their actual feedback. In regards to the partial compensation typically found in previous FAF studies, Larson et al. (2000) suggested that a complete model with this negative closed-loop feedback system would include a filter with a limiting nonlinearity preventing responses to over 50 cents. Thus, participant's compensatory responses will not exceed approximately half a semitone during FAF studies, which has mostly been the case in the literature examining speech production (Burnett et al., 1997, 1998; Hain et al., 2000; Natke et al., 2003; Larson, 1998). Accordingly, when pitch deviations occur during mid-utterance the pitch-shift reflex attempts to stabilize voice F_0 rather than adjust the pitch to match a target. Interestingly, when the perturbation was presented at utterance onset the compensatory responses found by Hawco and Jones (2009) were larger than those obtained when the perturbation was presented at mid-utterance. Thus, the mechanisms for vocal control responsible for compensatory responses appear to be modulated by the onset of the perturbation, whether it coincides with vocal production or occurs during the utterance.

It has also been proposed that compensatory responses occur as a result of servomechanistic processes. For example, servomechanistic explanations put forward that a comparator actively seeks for discrepancies between the motor commands with the actual sensory feedback received during vocal production.

Essentially, this relates to an 'efference' copy, or the motor plan, being compared to the actual feedback to determine whether an error exists. If an error is detected, the system will attempt to compensate for the perceived mismatch by altering the voice F_0 in the opposite direction of the perturbation. However, a conflict with this account exists, as average speech rates occur too quickly for auditory feedback to be reliably utilized in detecting and compensating for a perceived error prior to the next segment being produced (Bordon, 1979). As a consequence, a more likely explanation is that just before and immediately after vocalization, vocal fold stiffness and laryngeal structures appear to be responding in an open-loop fashion (Watts, Murphy, & Barnes-Burroughs, 2003).

On the other hand, Jones and Munhall (2000) argue that vocal productions are governed by both closed and open-loop control. For instance, when learning to produce a target note participants may have a stronger reliance on auditory feedback during the initial acquisition stages. This would assist with the formation of sensorimotor representations that will guide future vocal productions. This idea is based on the premise that the feedback one receives while learning to perform a specific task creates an integrated sensorimotor representation that is directly related to the task (Proteau Marteniuk, Girouard & Dugas, 1987; Proteau, Marteniuk & Levesque, 1992). Thus, over time, it is thought that an internal representation is formed, or 'internal model'. Internal models are thought to exist as neural maps of skilled movements that store the relationships between the motor commands, environment and sensory feedback for their production (Desmurget & Grafton, 2000; Flanagan & Wing, 1993;

Shadmehr & Mussa-Ivaldi, 1994). In regards to speech and singing, it is thought that the reliance on auditory feedback over time may be reduced in favour of an increased reliance on the motor plan for vocalization. Thus, if there exists an internal model for singing one would assume that it would be more entrenched in singers, by virtue of their extensive training, than in nonsingers.

Internal Models for Vocal Control: Musical Training

Jones and Keough (2008) conducted a study to investigate whether trained singers, given their extensive training, would rely more on an internal representation than nonsingers while singing. Participants produced a musical target (392 Hz or 349 Hz, G4 or F4 respectively) while receiving unaltered auditory feedback (during a baseline and test phase) or FAF (during a training phase), where auditory feedback was shifted down 1 semitone (100 cents). It was hypothesized that participants' would compensate for the FAF by altering their voice F_0 in the opposite direction of the perturbation. Moreover, both groups were predicted to recalibrate their internal models for F_0 control during FAF trials. Once this internal model is recalibrated, it was hypothesized that when subjects received normal unaltered feedback, after a period of receiving manipulated feedback, that aftereffects would persist for both singers and nonsingers. Thus, it was thought that the F_0 values obtained during testing would be significantly higher than participants baseline F_0 values. However, if trained singers possess a stronger reliance on an internal model to guide F_0 productions it was hypothesized that they should exhibit a smaller magnitude of compensation than

nonsingers when exposed to frequency-altered feedback. If singers relied more on an internal representation to regulate voice F_0 then following FAF aftereffects should be more apparent and persistent throughout testing than the aftereffects observed for nonsingers. Additionally, aftereffects should generalize to a stronger degree in singers than nonsingers when required to emulate a different note (F4) than the one they received during training.

In accord with the predictions made by Jones and Keough (2008), it was determined that when initially exposed to FAF that nonsingers compensated to a greater degree than singers. This difference persisted across nearly every block of the FAF trials during the experimental session. Of particular interest were the F_0 values collected for singers and nonsingers during the test trials. Singers' F_0 values were higher than their baseline F_0 values, whereas there were no differences observed for nonsingers' F_0 values during the baseline and test trials. Thus, aftereffects were only present for singers and these effects were found even when singers were required to produce a different note than the one they sang during training. As a result, these data imply that singers rely more on an internal representation to regulate their F_0 productions while singing.

My Dissertation: An Overview

The purpose of this dissertation is to further our understanding of the role of auditory feedback on internal representations controlling voice F_0 while singing. Moreover, I am interested in examining whether formal music training results in any differences between singers and nonsingers ability to use auditory

feedback to control the pitch of their voice. As a consequence, Chapter 1 will examine the sensitivity of trained singers and nonsingers' internal representations to subtle changes in auditory feedback. Unlike the previous study in our laboratory (Jones & Keough, 2008) where participants were presented with large (1 semitone; 100 cents) pitch shift manipulations, Chapter 1 had participants produce musical targets while receiving auditory feedback that was manipulated in +/- 2 cent increments across trials up to +/-100 cents (1 semitone). The purpose of this study was to address when (at what trial) participants began compensating for small changes in auditory feedback. It was thought that determining when participants, on average, compensated for FAF might elucidate their compensatory threshold to these subtle manipulations. Moreover, given the extensive training singers possess it was hypothesized that they should be more sensitive to subtle changes in their auditory feedback, resulting in compensatory responses occurring earlier than in nonsingers.

Additionally, we were interested in determining whether sensorimotor adaptation was observable during FAF trials. That is, could we observe changes in F_0 production over time when participants compensated for small pitch shift manipulations? In order to accomplish this we examined singers and nonsingers median F_0 values within 50 ms of vocal onset. If participants exhibited sensorimotor adaptation, then their F_0 values should progressively decrease or increase when feedback was gradually shifted upwards or downwards, respectively. Indeed, this should identify whether internal models are continuously updated and whether singers exhibit a stronger reliance on internal

models to control voice F_0 . Moreover, it was hypothesized that modifications in singers' internal representations would result in aftereffects that would generalize to other notes that were not manipulated in pitch to a greater degree than in nonsingers.

Chapter 2 was designed to investigate the sensitivity of trained singers' internal representations to altered auditory feedback and determine whether participants can develop and switch between multiple internal representations based on a contextual cue (musical notes). This experiment was largely based on the evidence presented in arm reaching studies, which suggests that the motor system can acquire multiple internal models when presented with a contextual cue (e.g., colour). The idea that multiple internal models may exist could potentially explain how an organism is capable of adapting to different perturbations in diverse contexts. Participants were presented with target notes (A4, G4, F4) on each trial prior to initiating vocal productions. Although participants were presented with target notes in a sequential order, they were only required to reproduce them one at a time when prompted by a cue (the target note). During FAF trials, auditory feedback for the first note in the sequence (A4 or F4) was gradually manipulated (over 25 trials) upward in pitch (in 4 cent increments to 100 cents; 100 cents = 1 semitone). Auditory feedback for the second target (G4) in the sequence remained unaltered during all trials. Thus, G4 was used as a control note to examine the sensory motor representation of an unaltered target while compensating for FAF. Auditory

feedback for the third note in the sequence (F4 or A4) was gradually manipulated (over 25 trials) downward in pitch (in -4 cent increments to -100 cents).

We hypothesized that if participants can form multiple internal representations for F_0 control, then they will exhibit a pattern of compensation that is unique to each target. Also, if trained singers initiated vocal productions at levels close to the target frequency then they should gradually adjust how they initiate voice F_0 (within 50 ms). That is, participants should initiate their vocal pitch at progressively lower and higher frequencies when presented with upward and downward pitch shifts within an experimental session, respectively. Furthermore, we predicted that if participants rely on multiple internal models, then F_0 values for the unaltered pitch target should remain unchanged even while recalibration has occurred for altered pitch targets.

In regards to Chapter 2, a question was raised as to whether singers will constantly update an internal model that was unique to each target note or will they rely on the predictability of the sequence to regulate F_0 control (each note corresponded to a specific manipulation)? In order to assess this, we decided to run additional trained singers using the same method, however we had participants emulate one target note (G4) while receiving the various FAF manipulations. It was predicted that if singers were relying on the predictability of the sequence then they should show a pattern of adaptation that is target specific. This would suggest that sensorimotor adaptation was due to the predictable nature of the task (anticipation of FAF) and not necessarily due to the detection of an error in the notes themselves. On the other hand, if singers are

constantly updating their internal models, and are not solely relying on the predictability of the sequence, they should not adapt during FAF trials in Experiment 2. Instead, producing the same note repeatedly with varying manipulations in auditory feedback should result in consistent F_0 values during vocal initiation (within 50 ms) with compensatory responses occurring later in production.

Chapter 3 had singers and nonsingers produce target notes at specific frequencies with different instructions while receiving subtle and large modifications in auditory feedback. Participants were informed that their auditory feedback was manipulated in pitch and they were instructed to either (1) 'compensate' for these changes by altering their voice F_0 in the opposite direction of the perturbation or (2) to 'ignore' their auditory feedback maintain their voice F_0 at frequencies similar to when their feedback was unaltered. The purpose is to investigate (i) whether task instructions influence compensatory responding (F_0 values within 1500 ms of vocal onset) and sensorimotor adaptation (F_0 values within 50 ms of vocal onset) that is typically observed while singing with unaltered and frequency-altered feedback. Moreover, (ii) collecting data from singers and nonsingers will help identify whether musical training influences acoustic-motor control when instructed to compensate or ignore auditory feedback manipulations. Regardless of whether the pitch manipulations are small (-2 cent increments down to -100 cents, where 100 cents = 1 semitone) or large (-100 cents), if both singers and nonsingers can suppress or eliminate compensatory responses and sensorimotor adaptation then it will suggest that

these responses are, to a certain degree, under volitional control. Moreover, it would also imply that auditory feedback, as well as musical training, is not entirely responsible for the continued maintenance of the motor control of singing. Conversely, if similar patterns of compensatory responding and sensorimotor adaptation are observed, then it will suggest that these responses are automatic and that auditory feedback has an important role in vocal motor control while singing.

It is hypothesized that both singers and nonsingers will be unable to ignore subtle shifts (-2 cent increments to -100 cents; 'ramp condition') in FAF. As a result, it is expected that participants will exhibit similar patterns of compensatory responding and sensorimotor adaptation during the ramp condition. When auditory feedback is suddenly shifted to -100 cents ('constant condition'), it is hypothesized that nonsingers will immediately compensate by increasing their voice F_0 in the opposite direction of the manipulation. Over time, it is believed that nonsingers will exhibit sensorimotor adaptation while compensating for the FAF. On the other hand, trained singers may be able to ignore large changes in auditory feedback when instructed to do so, which would be congruent with the results of Zarate and Zatorre (2008). Indeed, if this occurs then vocal productions should be similar to those when auditory feedback was unaltered (baseline). However, it should be noted that participants in Zarate and Zatorre (2008) received FAF following vocal onset (the manipulations occurred between 1000-1500 ms into production), whereas FAF coincided with vocal onset in the current study. It is possible that the suppression of compensatory

responding exhibited by trained singers in Zarate and Zatorre resulted from their reliance on the auditory feedback they received prior to the presentation of the FAF. Thus, trained singers may not be able to suppress compensatory responding or sensorimotor adaptation when presented with large changes in auditory feedback that coincide with vocal onset.

Chapter 1

Running head: ONSET OF VOCAL PRODUCTIONS

The Sensitivity of Auditory-Motor Representations to Subtle Changes in Auditory
Feedback while Singing.

Dwayne Keough¹ & Jeffery A. Jones ^{*1,2}

¹ Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

² Centre for Cognitive Neuroscience
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

*Author to whom correspondence should be addressed:

Jeffery A. Jones
Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

E-mail: jjones@wlu.ca
Tel: 519-884-0710 ext: 2992
Fax: 519-746-7605

Abstract

Singing requires accurate control of the fundamental frequency (F_0) of the voice. This study examined trained singers and untrained singers (nonsingers) sensitivity to subtle manipulations in auditory feedback, and the subsequent effect on the mapping between F_0 feedback and vocal control. Participants produced the consonant-vowel /ta/ while receiving auditory feedback that was shifted up and down in frequency. Results showed that singers and nonsingers compensated to a similar degree when presented with frequency-altered feedback (FAF), however, singers' F_0 values were consistently closer to the intended pitch target. Moreover, singers initiated their compensatory responses when auditory feedback was shifted up or down 6 cents or more, compared to nonsingers who began compensating when feedback was shifted up 26 cents and down 22 cents. Additionally, examination of the first 50 ms of vocalization indicated that participants commenced subsequent vocal utterances, during FAF, near the F_0 value on previous shift trials. Interestingly, nonsingers commenced F_0 productions below the pitch target and increased their F_0 until they matched the note. We conclude that singers and nonsingers rely on an internal model to regulate voice F_0 , but singers' models appear to be more sensitive in response to subtle discrepancies in auditory feedback.

Keywords: frequency-altered feedback, auditory feedback, internal model, fundamental frequency, F_0 , pitch, singing, singer.

The Sensitivity of Auditory-Motor Representations to Subtle Changes in Auditory Feedback while Singing.

The role of auditory feedback during vocalization has been a topic of much scientific inquiry. Both speech and singing are debatably the most complex motor actions humans are capable of producing. In order to produce a word, or to sing a musical note, one must possess strict control over respiratory muscles in addition to control over intrinsic and extrinsic laryngeal muscles. This control is achieved by an intricate network of cortical and brainstem areas dependent upon auditory (Sapir, McClean, & Larson, 1983; Larson, Altman, Liu, & Hain, 2008) and proprioceptive (Kirchner & Wyke, 1965; Wyke, 1974; Yoshida et al., 1989) reflex mechanisms. However, nonreflexive systems that utilize auditory feedback also contribute greatly to the development of speech in children (e.g., Oller & Eilers, 1988).

Postlingually, auditory feedback also appears to be monitored and used during ongoing speech. For instance, delaying auditory feedback results in disruptions in the quality of vocal productions (Yates, 1963). As well, altering auditory feedback typically elicits compensatory responses in speakers' ongoing vocal productions. Increases in masking noise and decreases in side-tone amplitude (Bauer, Mittal, Larson, & Hain, 2006; Lane & Tranel, 1971), shifts in formant frequencies (Houde & Jordan, 1998; Purcell & Munhall, 2006b), and the fundamental frequency (F_0) (Burnett, Senner, & Larson, 1997; Elman, 1981; Jones & Munhall, 2000; Kawahara, 1998) all cause vocal responses that oppose

the respective manipulations. The purpose of the present study is to further our understanding of the role of auditory feedback for the control of F_0 .

Frequency-altered feedback (FAF) is one paradigm that has been used to examine the importance of auditory feedback for the regulation of voice F_0 . In a typical FAF study, participants receive auditory feedback that has been manipulated in pitch to be higher or lower than their actual vocal productions. Responses to these perturbations generally result in decreases or increases in F_0 to upward or downward pitch shifts, respectively. Elman (1981) was among the first to investigate the behavioural effects of manipulating speakers' auditory feedback regarding F_0 . Since that time, the majority of FAF studies (e.g., Bauer et al., 2006; Burnett et al., 1997; Burnett et al., 1998; Burnett & Larson, 2002; Elman, 1981; Houde & Jordan, 1998; Jones & Munhall, 2000, 2002, 2005; Natke et al., 2003; Toyomura et al., 2007) have required participants to produce a relative target and not to match a specific frequency, as in singing. In those studies, most participants compensated for the discrepancy detected between perception and production, however, not perfectly (e.g., Burnett et al., 1997; Donath et al., 2002; Larson, 1998; Larson, Burnett, Kiran, & Hain, 2000; Natke et al., 2003). The response magnitudes previously reported for manipulations varying from +/- 100 to 600 cents (cents: a logarithmic unit used to measure small intervals between different frequencies, where 100 cents is equal to 1 semitone) ranged from approximately 15-65 cents with an average response magnitude of around 50 cents (regardless of the direction of manipulation).

Singing offers researchers a unique window into the study of F_0 control. Singing involves producing a succession of musical sounds at a particular (absolute) frequency (e.g., 440 Hz, A4). Thus, singers must maintain accurate vocal control over their F_0 to hit the desired notes. Sundberg (1987, p. 177) reported that if deviations occur from the intended target with the actual production, trained singers are able to compensate and match the tone (A4, 440 Hz) with an accuracy of less than 1 Hz. However, few studies have directly examined the role of auditory feedback while singing (Burnett et al., 1997; Jones & Keough, 2008; Natke, Donath, & Kalveram, 2003; Zarate & Zatorre, 2005, 2008). The results obtained during singing studies that have used FAF (e.g., Burnett et al., 1997; Natke, Donath, & Kalveram, 2003) indicate that the compensation is comparable to those obtained during speech production. The compensatory responses observed in trials of altered feedback indicate that voice production is regulated in a closed-loop fashion (Fairbanks, 1954; Larson et al., 2000; Lee, 1950). However, laryngeal structures and vocal fold stiffness are set in place prior to vocal onset (e.g., Watts, Murphy, & Barnes-Burroughs, 2003), suggesting a role for open-loop motor planning.

In the case of singing, when learning to produce a target note, stronger reliance on auditory feedback during the initial acquisition stages may be required in order to establish sensorimotor representations that will guide future vocal productions. This idea is based on the premise that the feedback one receives while learning

to perform a specific task creates an integrated sensorimotor representation that is directly related to the task (Proteau Marteniuk, Girouard & Dugas, 1987; Proteau, Marteniuk & Levesque, 1992). For instance, Finny and Palmer (2003) demonstrated that the availability of auditory feedback while pianists learned a piece of music significantly improved their ability to play the piece from memory during later recall. However when pianists were asked to perform well rehearsed musical sequences from memory, the removal of auditory feedback did not affect musical performance (Finny & Palmer, 2003). Thus, over time, it is hypothesized that an internal representation is formed, or 'internal model' that stores the relationships between the motor commands, environment and sensory feedback for their production (Proteau et al., 1987; Proteau et al., 1992). As a result, pianists may be capable of performing a well rehearsed piece from memory without auditory feedback because they are relying on internal motor commands that correspond to the musical composition. The same reliance on an internal motor plan may also exist for singing a well-rehearsed piece. For instance, it is possible that vocal productions are initiated based on an increased reliance on the motor plan established for vocalization.

Jones and Keough (2008) investigated whether trained singers, given their extensive training, rely more on a well-established internal representation than nonsingers while singing. Indeed, identifying differences in F_0 control while speakers receive subtle manipulations in auditory feedback should elucidate the sensitivity of the underlying mechanisms regulating F_0 control. After training

participants to emulate a musical target (392 Hz, G4) while receiving auditory feedback shifted down by one semitone (100 cents), the authors found that nonsingers initially compensated to a greater degree than singers when exposed to FAF (Jones & Keough, 2008). However, after this brief exposure to FAF, singers' F_0 values were higher than their baseline F_0 values when they heard their feedback returned to normal, whereas there were no differences between F_0 values during the baseline and test trials for nonsingers. These effects persisted when singers were required to produce a different note other than the one they sang during the altered feedback trials. These results imply that singers update their internal forward model based on sensory feedback whereas nonsingers rely more on auditory feedback to regulate their F_0 .

The present study was designed to investigate the sensitivity of singers and nonsingers' internal representations to subthreshold (2 cent increments) manipulations in auditory feedback. The data obtained by Jones and Keough (2008) suggested that singers do not compensate for larger discrepancies between perception and production and instead rely on their internal models. Nonsingers fully compensated immediately after exposure to altered feedback of 100 cents whereas singers required several trials to modify their F_0 (Jones & Keough, 2008). Moreover, Zarate and Zatorre (2005, 2008) found that singers were able to successfully ignore altered auditory feedback (by presumably relying on an internal model) they received and continue to produce the targets at the desired frequency when instructed to do so. Conversely, nonsingers' vocal

productions suggested that they were unable to ignore the altered feedback and as a result they adjusted their F_0 to compensate for the manipulation (Zarate & Zatorre, 2005, 2008).

Jones and Keough (2008) found that singers initially responded less to large (100 cent) shifts in auditory feedback than nonsingers, who exhibited near perfect levels of compensation almost immediately. However, it remains unknown whether similar F_0 values for singers and nonsingers would be obtained using small (2 cent increments up to 1 semitone, 100 cents) manipulations in auditory feedback. Given the paucity of data, we hypothesized that singers and nonsingers would compensate to a similar degree to the altered-feedback, but that singers, due to their extensive training and experience, would compensate for the discrepancy in perception and production more efficiently than nonsingers. That is, the mechanisms that regulate F_0 control in singers would be more sensitive to subtle changes in FAF. As a result, singers will reproduce the target notes more accurately than nonsingers while receiving FAF. Of particular interest was determining when singers and nonsingers initiate compensatory responses to small manipulations in auditory feedback. Determining when both groups initiate compensatory responses may also provide an index of how sensitive the underlying mechanisms of the internal representations are to perturbations. Recently, the just-noticeable difference of fundamental frequency to digitally synthetic stimuli has been shown to be between 4-98 cents (Pape & Mooshammer, 2006). Furthermore, Loui and colleagues (2008) found that

participants without speech or hearing disorders, and with no formal music training, had a perceptual and production threshold (approximately 2.0 & 2.5 Hz, respectively) that was significantly better than the threshold for amusic participants (approximately 36.2 & 12.3 Hz, respectively). We hypothesized that differences singers would initiate compensatory responses to altered feedback earlier than nonsingers due to their extensive vocal experience achieving pitch targets.

Previous studies using the FAF paradigm have found evidence for sensorimotor adaptation by measuring the mean F_0 after feedback was returned to normal (Jones & Munhall, 2000, 2002, 2005; Jones & Keough, 2008). These aftereffects can be observed following exposure to a single trial of FAF (Natke, Donath and Kalveram, 2003) in nonsingers. To track sensitivity to FAF in this experiment, we assessed sensorimotor adaptation by measuring F_0 at vocal onset while exposing participants to FAF that increases or decreases in frequency by 2 cents on each successive trial. Using this measure, aftereffects will be evident when voice F_0 values start at or near F_0 values observed at the end of the previous utterance. If internal models are continuously updated, then onset of vocalization should progressively become flatter and sharper when participants receive FAF that is either increasing or decreasing, respectively. Thus, when participants receive unaltered auditory feedback at the end of the experiment, adaptation will be demonstrated if their F_0 values begin near their F_0 value during the final FAF trial. As a result, if singers rely more on an internal representation to control voice F_0 then aftereffects should be more pronounced in singers than in nonsingers.

Moreover, adaptation effects should generalize to a greater degree in singers than in nonsingers when asked to produce a different note (A4 & F4) with unaltered feedback after FAF trials.

Methods

Participants

Twenty Wilfrid Laurier University students (all women) whose native tongue was North American English participated. Although there is no evidence to suggest that a gender difference exists in response to FAF, men were excluded so that all participants could adequately sing the same target notes. Of the twenty participants, 10 were trained singers (mean musical training was approximately 12 years) recruited from the faculty of music at Wilfrid Laurier University. The remaining 10 participants were considered to be nonsingers, as none possessed any form of vocal training or ongoing participation in formal singing. All participants received financial compensation and gave informed consent. The Wilfrid Laurier University Research Ethics Committee approved the procedures.

Apparatus

Participant Recording Sessions. Participants were situated in a double-walled sound attenuated booth (Industrial Acoustic Company, Model 1601-01) and were

fitted with headphones (Sennheiser HD 280 Pro) and a condenser microphone (Countryman Isomax E6 Omnidirectional Microphone), which was approximately 3 cm from their mouth. Multitalker babble noise (Auditec, St. Louis, MO) was presented at 80 dB SPL to limit natural acoustic feedback. The target notes consisted of a female voice singing the consonant-vowel /ta/, that was presented at either 349, 392 or 440 Hz, respectively. Microphone signals were sent to a signal processor (VoiceOne 2.0, TC Helicon) that manipulated auditory feedback. The altered feedback was then mixed (Mackie ONYX 1640) with the multi-talker babble and subsequently sent to the participant. Vocal productions were digitized at 44.1 kHz for future analysis.

Target stimuli recording. The stimuli used were the same as in Jones and Keough (2008). A trained singer produced the respective targets, F4, G4, and A4, which were processed using the speech modification algorithm STRAIGHT (Speech Transformation and Representation using the Adaptive Interpolation of weighted spectrum; Kawahara, 1999) to ensure that each target was exactly 349, 392 or 440 Hz.

Procedure

Participants emulated a musical target over 210 trials, which were divided into 3 blocks of 70 trials. Each block consisted of 10 baseline, 50 shift and 10 test trials. On one block, participants reproduced the musical target G4 on all trials. Thus,

participants received unaltered auditory feedback during the 10 baseline trials, followed by 50 FAF trials (2 cent increments to 100 cents, 1 semitone) and then 10 more unaltered feedback trials. In two other blocks, participants reproduced either F4 or A4 for the 10 baseline and 10 test trials while singing G4 during the FAF trials. Thus, participants only sang G4 during FAF trials. The other target notes, A4 and F4, were used to test whether participants exhibited a transfer of adaptation when emulating notes that were never manipulated. Thus, the three blocks participants experienced were AGA, FGF and GGG, where the first letter denotes the baseline target (trials 1-10), the middle letter denotes the shifted target (trials 11-60) and the final letter denotes the test value (trials 61-70). The three blocks of trials were counterbalanced across participants.

On the first day of testing, participants produced the target on the 3 blocks (AGA, FGF and GGG) of 70 trials and received auditory feedback that was shifted either upward or downward in 2 cent increments up to 100 or -100 cents, respectively during the shift trials. Note that auditory feedback was shifted from the beginning of each utterance until the end of their vocal productions. On a subsequent day, participants produced the remaining 3 blocks of 70 trials while receiving altered feedback that was shifted in the opposite direction. Thus, participants were required to produce the target for a total of 6 blocks of 70 trials that were counterbalanced over two days. Individual trials commenced with the presentation of multitalker babble for 1000 ms followed by the target stimulus, which lasted 2000 ms in duration, and then the multitalker babble was presented

again for 4000 ms. The babble was presented prior to the target as a way to inform participants that the target note would be forthcoming. During the second presentation of the babble, participants were instructed to sing the target note as accurately as possible in pitch and duration. Trials were initiated and controlled by a computer and participants' vocal productions were recorded on a Macintosh G5 computer (Tracktion v1, Woodinville, WA). F_0 values for each vocal production were calculated, during offline analyses, using an autocorrelation algorithm included in the Praat program (Boersma, 2001). F_0 values were normalized to each target note (F4, G4 or A4) by calculating the appropriate cent values using the following formula:

$$\text{Cents} = 100 (12 \log_2 F/B)$$

Where F is the F_0 value in Hertz and B is frequency of the target pitch participants were instructed to sing (349, 392 or 440 Hz).

Results

The mean F_0 values for singers and nonsingers were calculated for each condition and are displayed in Figures 1a and 1b (grey lines), respectively. Data for one nonsinger was removed from the statistical analyses due to the fact that she exhibited poor F_0 control during testing (mean F_0 was 346 and 244 cents for the shifted up and down conditions, respectively). Only the first 1500 ms for each vocal production was analyzed because previous research has found

compensatory responding to FAF occurs between 130 to 500 ms after perturbation onset (Burnett et al., 1997; Burnett et al., 1998; Jones & Munhall, 2002). The F_0 values from the pitch shift up and down AGA, FGF and GGG blocks were analyzed together and were broken into five different blocks of trials within each condition: shift trials (11-20, 21-30, 31-40, 41-50, 51-60).

Furthermore, the first five trials and the last five trials of each block were averaged and divided into early and late phases, respectively. Thus, a repeated-measures ANOVA was carried out on the mean F_0 values with 2 (experience: singer and nonsinger) x 2 (pitch shift: up and down) x 5 (block) x 2 (phase: early and late) as factors. Newman-Keuls' test was used for post-hoc tests with an alpha level of .05 used for all statistical tests.

Also, we wanted to determine when singers' and nonsingers' mean F_0 values during altered feedback trials were significantly different than baseline F_0 values. This test would indicate when compensation occurred to the subtle manipulations in feedback. Furthermore, we examined the first test trial immediately after participants received altered feedback to determine if sensorimotor adaptation, in the form of aftereffects, occurred. In order to assess compensatory responses and aftereffects, multiple t-tests were performed on the mean F_0 value for the last five baseline trials compared with the mean F_0 values on altered feedback trials and the first test trial during shifted up and down conditions.

Additionally, the median F_0 value for the first 50 ms (see Figures 1a & 1b – black lines) of each utterance during the AGA, FGF and GGG conditions was calculated for both singers and nonsingers across shifted trials. Determining whether differences occur between the initial 50 ms of initial shift trials and the later shifted trials identified whether sensorimotor adaptation occurred during successive FAF trials. Furthermore, these median values were subtracted from the mean F_0 over the entire utterance to determine the difference between where participants initiated vocal pitch and where they maintained their F_0 values while singing. This test identified whether singers and nonsingers initiated F_0 productions at the desired target frequency or whether they performed a searching strategy, increasing or decreasing their F_0 to match the target. The median values for the first 50 ms and the median less the mean values were then categorized into five blocks of trials and further divided into two phases, in the same fashion as the mean F_0 values. Thus, two separate ANOVAs were performed on each of the aforementioned values (median & median minus mean) during the shifted up and down AGA, FGF and GGG conditions with 2 (experience: singer and nonsinger) x 2 (pitch shift: up and down) x 5 (block) x 2 (phase: early and late) as factors.

Finally, in order to elucidate when or if sensorimotor adaptation commences during FAF trials and whether aftereffects persist during test trials, multiple t-tests were conducted on the median F_0 values for the first 50 ms during both the shifted up and down conditions. The average of the median F_0 values for the last

five baseline trials were compared with the median F_0 values of all shift trials and the first test trial following exposure to FAF.

Results of the mean F_0 values during the shifted up and down conditions revealed a main effect of experience and pitch shift, $F(1, 55) = 12.59$, $p < .05$ and $F(1, 55) = 634.73$, $p < .05$, respectively. Nonsingers' mean F_0 values were found to be significantly flatter (lower than the pitch target of 392 Hz, G4) than the mean F_0 values of singers (see Figure 1a and 1b). Also, the mean F_0 values during the pitch shift up condition were found to be flatter than the F_0 values during the pitch shift down condition. A significant two-way interaction was found between pitch shift and block, $F(4, 220) = 423.40$, $p < .05$. Post-hoc analysis revealed that the first block of mean F_0 values during the pitch shift up condition and the pitch shift down condition were significantly different (progressively flatter and sharper during pitch shift up and down conditions, respectively) than all other blocks of pitch shift F_0 values ($p < .05$) (see Figure 1a and 1b). This pattern demonstrates that participants were compensating more during later trials to increasing and decreasing shifts in auditory feedback. Interestingly, this suggests that both singers and nonsingers initiate compensatory responses quite early to subtle changes in auditory feedback. Furthermore, the mean F_0 values obtained for each pitch shift block during the shift up condition were significantly flatter than the mean F_0 values for each block in the shift down condition ($p < .05$).

A two-way interaction between pitch shift and phase was also found to be significant, $F(1, 55) = 121.05$, $p < .05$. Post-hoc analysis indicated that the mean F_0 values during the first phase of the shift up condition and the shift down condition were significantly different than the mean F_0 values during the second phase of the shift up and down conditions, respectively ($p < .05$) (see Figure 1a and 1b). This implies that as pitch shifts progressively increased or decreased on shifted trials participants adjusted their F_0 to continue to produce the target notes, as late phase shift trials were larger than early phase shift trials. Also, the F_0 values during the first and second phase of the shift up condition were significantly different than the F_0 values of the first and second phase of the shift down condition ($p < .05$). No other significant main effects or interactions were observed.

Multiple t-tests (uncorrected) were carried out on the mean F_0 during the shifted up and down conditions for singers and nonsingers. When three consecutive significant differences were found during the shifted trials the first significant response from baseline was used to indicate the initial compensatory response. During the shift up trials, when the mean of the last five baseline trials was compared with the mean of the shift trials it was found that singers initiated compensatory responses on the third (6 cents) shift trial, $t(60) = -2.89$, $p < .05$. The same initial response was also observed during the shift down condition, $t(60) = 2.67$, $p < .05$. Moreover, singers F_0 values on shift up and down trials remained significantly different than their baseline F_0 values. Thus, singers F_0

values on shifted trials were significantly different than the average of the last five baseline trials on 48 of 50 trials during both shift up and down conditions.

On the other hand, during the shift up condition nonsingers were found to initiate compensatory responses on shift trial 13 (26 cents), $t(60) = -3.68$, $p < .05$, with the remaining F_0 values remaining significantly different than the baseline F_0 values. The only difference between the shift up and down condition for nonsingers was that they initiated compensatory responses two trials earlier during the down condition, at shift trial 11 (22 cents), $t(60) = 3.12$, $p < .05$. Thus, nonsingers' F_0 values were different than baseline on 41/50 and 40/50, during the shift up and down conditions respectively.

We were also interested in determining whether compensating for FAF would result in aftereffects (for the GGG condition) when auditory feedback was returned to normal and whether these aftereffects would generalize (for the AGA and FGF conditions) to a note other than the one participants received during testing. Thus, t-tests were carried out on the average of the mean F_0 values of the last five baseline trials for the AGA, FGF and GGG conditions and the first test trial (when auditory feedback was returned to normal following FAF trials) of each respective condition (see Figure 2a and 2 b). Nonsingers' average baseline F_0 values for the GGG trials were significantly different than the F_0 values for the initial test trial for both shift down and up conditions, $t(8) = -2.90$, $p < .05$ and $t(8) = 2.30$, $p < .05$ respectively. Moreover, singers baseline mean F_0 values for

both shift down and up conditions were also significantly different than their initial mean F_0 values during testing, $t(9) = -5.39$, $p < .05$ and $t(9) = 4.49$, $p < .05$, respectively. Thus, both singers and nonsingers exhibited aftereffects following FAF trials. That is, their mean F_0 values were above and below the target F_0 when auditory feedback was returned to normal following exposure to FAF.

When participants sang a different note following FAF trials, nonsingers' mean F_0 values were only significantly different than their average baseline F_0 values for the shifted up AGA condition, $t(8) = 2.58$, $p < .05$. On the other hand, singers' mean baseline F_0 values were significantly different than their mean test F_0 value for both the AGA and FGF shifted down condition, $t(9) = -2.35$, $p < .05$ and -3.72 , $p < .05$, respectively, and the AGA shifted up condition, $t(9) = 5.51$, $p < .05$. Overall, both groups exhibited aftereffects that generalized to at least one note other than the one produced during FAF trials. However, when the alpha level was corrected for multiple t-tests, nonsingers' mean F_0 baseline values failed to remain statistically significant when compared to the initial test trial for all conditions. Singers' aftereffects during GGG trials remained significantly different during both shifted down and up conditions, but the only generalization effect that remained significant was during the shifted up AGA condition.

The median 50 ms F_0 values obtained during testing revealed a significant main effect of experience and pitch shift condition, $F(1, 55) = 15.11$, $p < .05$ and $F(1, 55) = 57.72$, $p < .05$, respectively. Nonsingers' median F_0 values were determined to be significantly flatter than singers' median F_0 values. The median

F_0 values during the shift down condition were found to be significantly sharper than the median F_0 values obtained during the shift up condition. A two-way interaction between pitch shift and block was significant, $F(4, 220) = 45.58$, $p < .05$. Post-hoc testing indicated that the median F_0 values during the first block of shift trials on the shift up condition was significantly sharper than all remaining shift blocks ($p < .05$). Moreover, the median F_0 values during the first block of shift trials during the shift down condition were significantly flatter than shift blocks 3, 4 and 5 ($p < .05$). Also, the median F_0 values for all shift up blocks were significantly flatter than the median F_0 values for all shift down blocks ($p < .05$).

Additionally, there was a significant interaction between pitch shift and phase, $F(1, 55) = 7.39$, $p < .05$. Post-hoc testing revealed that there were no differences between the early and late phases of the shift up and down conditions, respectively ($p > .05$). However, the median F_0 values of the early and late phase of the shift down condition were significantly sharper than the median F_0 values of both the early and late phase of the shift up condition ($p < .05$). Finally, there was a significant 3-way interaction between experience, pitch shift and block, $F(4, 220) = 3.01$, $p < .05$. Post-hoc analysis revealed that nonsingers' first block of median F_0 values during the initial block of the shift down condition were significantly flatter than the median F_0 values during blocks 3, 4 and 5 ($p < .05$). Also, during the shift up condition nonsingers' median F_0 values during the initial shift block were significantly sharper than the median F_0 values on blocks 3, 4 and 5 ($p < .05$). Singers initial shift block median F_0 values for both the shift down

and up conditions were significantly sharper and flatter, respectively, than the median F_0 values on shift blocks 3, 4 and 5 ($p < .05$). Moreover, singers' median F_0 values on shift down trials were significantly sharper than all nonsinger median F_0 values on shift down trials ($p < .05$). Singers' median F_0 values on shift up trials were also found to be significantly sharper than all nonsinger median F_0 values on shift up trials ($p < .05$). No other significant main effects or interactions were observed.

Multiple t-tests (uncorrected) were also performed on the median values of the first 50 ms of each utterance. Similarly to the multiple t-tests reported previously, three consecutive significant differences were required prior to establishing when adaptation occurred. The last five baseline median F_0 values were averaged to establish a baseline for comparison. This value was then compared to all shifted and test trials. During the shifted up condition, singers median F_0 values were found to be different than baseline at shift trial 17 (34 cents), $t(60) = -2.59$, $p < .05$. In total, singers' median F_0 values were different than baseline on 33/50 shifted up trials. During the shifted down condition, singers exhibited adaptation responses at shift trial 13 (26 cents), $t(60) = 3.94$, $p < .05$. During the shifted down condition singers, median F_0 values were different than baseline on 35/50 trials. Interestingly, nonsingers' median F_0 values failed to significantly differ on three consecutive occasions during the shifted up or down conditions. However, nonsingers' median F_0 values did differ on two consecutive trials during both the shifted up and down conditions (but only once during each). Nonsingers' F_0

values differed at shift trial 38 (76 cents) and 39 (78 cents), $t(60) = -2.73$ and 2.65 , $p < .05$ during shifted up and down conditions, respectively. Furthermore, nonsingers' median F_0 values only varied from baseline on 12/50 and 4/50 during the shifted up and down conditions, respectively.

When the median F_0 values for the first 50 ms of the shifted down and up conditions were subtracted from the mean F_0 values for each respective condition, results indicated that there was a significant main effect of experience, $F(1, 55) = 12.52$, $p < .05$. Nonsingers' F_0 values were found to be significantly flatter than the F_0 values for singers. A main effect of pitch shift was also found to be significant, $F(1, 55) = 6.90$, $p < .05$. F_0 values during the pitch shift down condition were found to be significantly sharper than the F_0 values during the pitch shift up condition. Results identified a significant 2-way interaction between pitch shift and block, $F(4, 220) = 6.86$, $p < .05$. Post-hoc testing indicated that the only difference during the pitch shift down condition was that the initial block of F_0 values were significantly flatter than the last block of shift trials ($p < .05$).

Moreover, the initial shift block F_0 values during the shift down condition was significantly different than the F_0 values of shift blocks 4 and 5 of the shift up condition ($p < .05$). The initial block of shift trials during the shift up condition were determined to be significantly sharper than shift blocks 4 and 5 ($p < .05$), as well as the last block of trials during the shift down condition ($p < .05$). Finally, there was a significant 3-way interaction between experience, pitch shift and block, $F(4, 220) = 2.72$, $p < .05$. Post-hoc testing revealed that nonsingers' first

block of F_0 values during the shift down condition were significantly flatter than the F_0 values during the last block of shift trials ($p < .05$). Nonsingers' F_0 values during the initial block of shift up trials were significantly sharper than the F_0 values of blocks 4 and 5 ($p < .05$). Singers' F_0 values during the initial block of shift down trials were not significantly different than any other block of shift down trials ($p > .05$), nor were the F_0 values of the initial block of shift up trials different than any other block of shift up trials ($p > .05$). Furthermore, nonsingers' F_0 values during the shifted up and down conditions were all significantly flatter than singers' F_0 values ($p < .05$). No other significant main effects or interactions were observed.

Discussion

The purpose of this study was to investigate the sensitivity of the mechanisms that modify internal representations in singers and nonsingers when presented with subtle changes in auditory feedback while singing. In accord with our predictions, singers and nonsingers use auditory feedback to compensate for subtle manipulations in auditory feedback while singing. That is, both singers and nonsingers compensated for the altered feedback by increasing or decreasing their F_0 to upward and downward shifts in feedback. The main difference between the groups was that singers are immediately and consistently more accurate when they emulate the target notes. Nonsingers' F_0 values are consistently flat, regardless of the direction of the manipulation. An important

observation was that singers initiated compensatory responding to altered feedback on the third shift (6 cents) trial during both upward and downward altered conditions. On the other hand, nonsingers initiated compensatory responding at shift trials 13 (26 cents) and 11 (22 cents) during upward and downward manipulations, respectively.

Furthermore, we examined the first 50 ms of each vocal production in order to determine how participants initiate F_0 responses while singing. If participants adapt to the altered feedback by altering an internal representation that regulates their initial F_0 production, their initial F_0 production should be close to the F_0 values produced while compensating during the previous trial. Results showed that singers, and to a lesser degree nonsingers, incorporate the discrepancy between perception and production by starting subsequent utterances, during shifted trials, where they ended on the previous production. On the other hand, nonsingers appeared to search for the target note by starting below the auditory target of 0 cents and increasing their F_0 until they matched the note. Moreover, multiple t-tests failed to provide evidence to suggest that sensorimotor adaptation occurred in nonsingers during the altered feedback conditions. In the case of singers, multiple t-tests on the median 50 ms F_0 data revealed that sensorimotor adaptation occurred at 34 cents and 26 cents during shifted up and down conditions, respectively.

Previous research using the FAF paradigm has typically examined speech (Bauer et al., 2006; Burnett et al., 1997; Burnett et al., 1998; Burnett & Larson, 2002; Elman, 1981; Houde & Jordan, 1998; Jones & Munhall, 2000, 2002, 2005; Natke et al., 2003; Toyomura et al., 2007), with an emphasis on compensatory responding. Not surprisingly, the results are consistent with the data collected on singing (Burnett et al., 1997; Jones & Keough, 2008; Natke, Donath, & Kalveram, 2003, Zarate & Zatorre, 2005, 2008), in that participants compensate for manipulations in auditory feedback by increasing or decreasing their F_0 in the opposite direction of the perturbation. The data obtained from the current experiment is in accord with this observation, as both singers and nonsingers altered their F_0 to similar degrees while receiving altered feedback.

It has been argued that the aftereffects observed after exposure to FAF are a result of modifying an acoustic-motor representation (Jones & Munhall, 2000, 2002; Jones & Keough, 2008; Natke, Donath & Kalveram, 2003). However, sensorimotor adaptation has not been extensively examined in singers. When Jones and Keough (2008) returned trained singers auditory feedback to normal, following FAF trials, their F_0 values never returned to baseline levels. Thus it appears that brief exposure to altered feedback resulted in a remapping of an internal representation (Jones & Keough, 2008). In the current study, we also found a global remapping to subtle changes in auditory feedback. Uncorrected multiple t-tests revealed that aftereffects existed for both singers and nonsingers. That is, when auditory feedback was suddenly returned to normal, following

exposure to FAF, singers and nonsingers mean test trial F_0 values were significantly higher and lower (following shifted down and up feedback, respectively) than baseline. Also, these aftereffects generalized to a different note other than the one they produced during FAF trials. For instance, singers' mean F_0 values for the target note A4 during shift up and down conditions and F4 following shift down trials was significantly different than the average of the last five baseline trials. Nonsingers' test F_0 values were only different than baseline for the target note A4 during the shift up feedback condition. Thus, participants modified an internal representation that regulates F_0 control. As a consequence, the pattern of responding observed in the current study is similar to the pattern of responding that has been found in previous FAF studies (Jones & Munhall, 2000, 2002; Jones & Keough, 2008). However, when the alpha level was corrected for the multiple t-tests, only the aftereffects observed for singers and the generalization for the target note A4 during the shifted up condition remained significant.

Previous studies have identified aftereffects following FAF, but have only examined adaptation after a period of training, at the end of an experiment (Jones & Munhall, 2000, 2002; Jones & Keough, 2008). For this study, we tracked F_0 values early in each utterance during the training period to determine whether singers and nonsingers continuously adjust their auditory-motor mapping in response to changes in feedback. Singers were found to emulate the target, on average, more accurately than nonsingers. That is, they initiated their

vocal productions near the F_0 frequencies that were required to compensate for the FAF experienced in previous trials. This sensorimotor adaptation in singers occurred when the discrepancy between expected F_0 output and auditory feedback was approximately 30 cents. Thus, subtle discrepancies in feedback can be accounted for by the mechanisms that support singers auditory-motor mapping. On the other hand, nonsingers consistently started their productions below the target (flat). Their F_0 values began flat and were adjusted until they reached the note, as best as they could. This difference between singers and nonsingers is consistent with the data obtained in Jones and Keough (2008).

One difference found in previous work investigating speech and singing using FAF was the level of compensation observed. Typically, it has been reported that during altered feedback compensatory responses do not exceed 65 cents for shifts up to 600 cents. As a consequence, Larson and colleagues (2000) have suggested these responses are a result of closed-loop control, and propose a model that includes a filter with a limiting nonlinearity that prevents responses greater than 50 cents. However the bulk of previous research has focused on speech. Jones and Keough (2008) reported that nonsingers fully compensated (100 cents) for altered feedback almost immediately and although singers initially showed partial compensation (approximately 65 cents) they eventually altered their F_0 values to accommodate for the altered feedback. Data from the current study revealed that singers and nonsingers exhibited near perfect levels of compensation during the shifted up and down conditions. Thus, it appears that

auditory feedback may be used in a task dependent manner such that when achieving a particular pitch target is important, as in singing, auditory feedback guides production.

Although singers and nonsingers eventually compensated for the FAF to the same degree, when each group altered their productions based on the auditory feedback differed. Singers began compensating when they detect feedback errors as large as 6 cents (1.36 Hz) during both the shift up and down conditions. This value is consistent with Sundberg's (1987) finding that trained singers can correct for production errors with an accuracy of less than 1 Hz from an intended pitch target (A4, 440 Hz). On the other hand, nonsingers initiated compensatory responses at approximately 24 cents (5.47 Hz). Data from both groups fall within the just-noticeable difference range found by Pape and Mooshammer (2006) of F_0 contours for digitally synthetic stimuli. Moreover, nonsingers' values are also very similar to the average threshold reported by Hafke (2008), who found that pitch shift changes were not reliably identified when they were below 26 cents. However, Loui and colleagues (2008) found that control participants psychophysical thresholds of perception and production were around 2.0 and 2.5 Hz, respectively. Although we never examined the perceptual aspect directly, the data we obtained for singers and nonsingers is relatively consistent with the production threshold findings of Loui et al. (2008).

Based on the current findings, we believe that singers, due to their extensive training and experience, are more capable of compensating for subtle manipulations of auditory feedback earlier and more efficiently than nonsingers. Moreover, singers more readily alter their internal representations to prevent the occurrence of these errors in subsequent utterances. On the other hand, when singers detect large incongruities between perception and production they rely more on their internal model to produce the target (Jones & Keough, 2008). The F_0 control system may deem certain feedback discrepancies to be too large to be internally generated. Zarate and Zatorre (2005, 2008) found similar results when singers and nonsingers were exposed to FAF. Zarate and Zatorre (2005, 2008) asked participants to ignore the feedback and continually reproduce the target as accurately as possible. The pattern of behavioural results suggests that singers could successfully ignore their altered (200 cents) feedback and continue to produce the target notes accurately. Interestingly, nonsingers were unable to ignore the feedback as their F_0 responses indicated partial compensation. These findings suggest that singers' internal models are flexible in nature; the models can be adjusted for small errors (act on auditory feedback) and can ignore large errors by relying on the existing motor representation.

Given that behavioral differences exist between singers and nonsingers during FAF, differences in neural activity between the two groups may be expected. Zarate and Zatorre (2005, 2008) found that despite differences in vocal production accuracy during normal feedback conditions, both singers and

nonsingers exhibited similar functional networks for singing. These areas included the bilateral auditory cortices, bilateral primary motor cortices, the supplementary motor area (SMA), the anterior cingulate cortex (ACC), thalamus, insula, and cerebellum. These results are consistent with a previous study carried out in the Zatorre laboratory (Perry et al., 1999) and by Toyomura and colleagues (2007). However, when exposed to FAF and asked to ignore the FAF, singers showed enhanced activation in the inferior parietal lobule (IPL), superior temporal gyrus (STG), superior temporal sulcus (STS), and right insula (Zarate & Zatorre, 2005). On the other hand, enhanced activation in the ACC, STS, insula, putamen, pre-SMA, and IPL was observed in singers when they were directly asked to compensate for the FAF (Zarate & Zatorre, 2005). The authors conclude that the additional activation of the STG and the STS in singers during FAF conditions are suggestive of an increased perceptual analysis of the incoming signal (Zarate & Zatorre, 2005). In addition, the authors point out that an increased activation in the putamen of singers during both ignore and compensate conditions suggests that singers are relying on well-defined internal representations to sing the targets while receiving FAF (Zarate & Zatorre, 2008).

Figure Caption

Figure 1 a

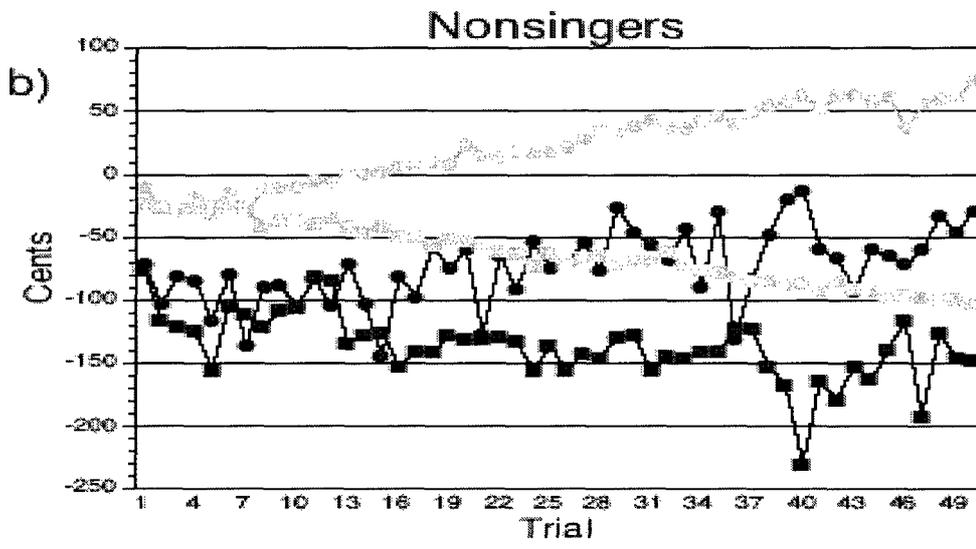
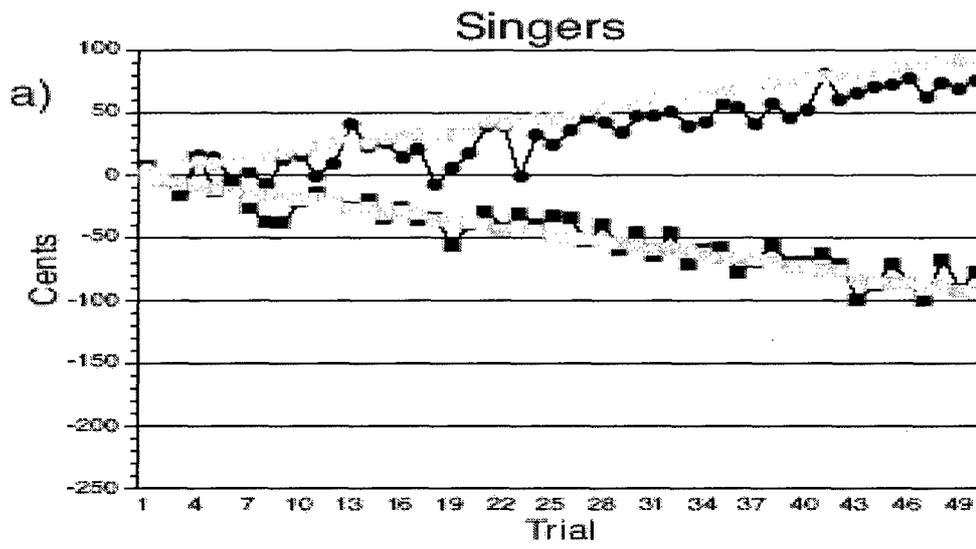
The mean and median 50 ms F_0 values for singers' during frequency-altered feedback trials when required to emulate the musical target note G4 (392 Hz). The grey circles represent singers' mean F_0 values and the black circles represent singers' median F_0 values during the initial 50 ms of vocal onset when auditory feedback was shifted downward. The grey squares represent singers' mean F_0 values and the black squares represent singers' median F_0 values during the initial 50 ms of vocal onset when auditory feedback was shifted upward.

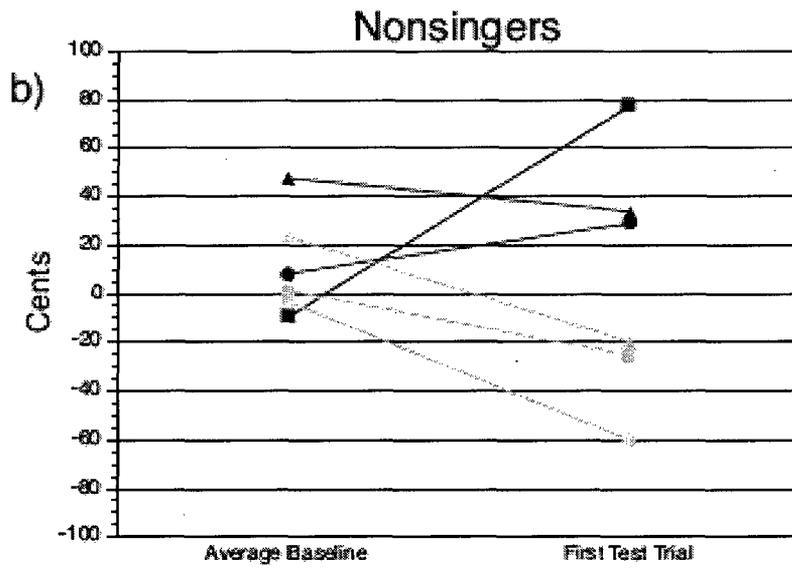
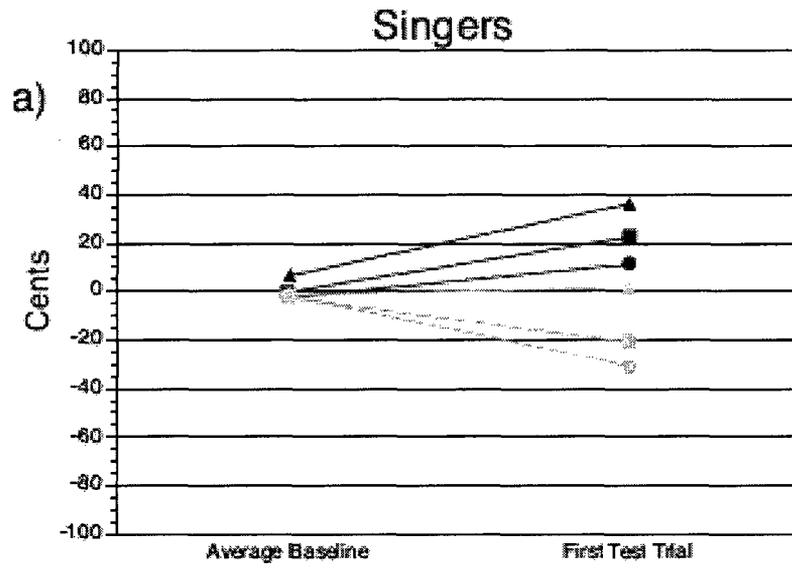
Figure 1 b

The mean and median 50 ms F_0 values for nonsingers during frequency-altered feedback trials when required to emulate the musical target note G4 (392 Hz). The grey circles represent nonsingers' mean F_0 values and the black circles represent nonsingers' median F_0 values during the initial 50 ms of vocal onset when auditory feedback was shifted downward. The grey squares represent nonsingers' mean F_0 values and the black squares represent nonsingers' median F_0 values during the initial 50 ms of vocal onset when auditory feedback was shifted upward.

Figure 2 a & b

The mean F_0 of utterances produced by (a) singers and (b) nonsingers, averaged over the last five base line trials prior to receiving FAF, and the mean F_0 for the first test trial utterance following FAF. The triangles represent F_0 values for the target note F4 349 Hz, the squares represent F_0 values for the target note G4 392 Hz, and the circles represent F_0 values for the target note A4 440 Hz. Gray symbols represent F_0 values obtained during the upward shift condition, and the black symbols represent F_0 values obtained during the downward shift condition.





Chapter 2

Contextual Cuing Contributes to the Independent Modification of Multiple Internal
Representations for Vocal Control

Dwayne Keough¹ and Jeffery A. Jones ^{*1,2}

¹ Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

² Centre for Cognitive Neuroscience
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

Running head: MULTIPLE INTERNAL REPRESENTATIONS FOR VOCAL
CONTROL

*Author to whom correspondence should be addressed:

Jeffery A. Jones
Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada
E-mail: jjones@wlu.ca
Tel: 519-884-0710 ext: 2992
Fax: 519-746-7605

Abstract

Research on the control of visually guided limb movements indicates that the brain learns and continuously updates an internal model that maps the relationship between motor commands and sensory feedback. A growing body of work suggests that vocal control is also supported by an internal model relating motor commands to auditory feedback. There is some evidence from arm reaching studies that the motor system can acquire multiple internal models, when provided with a contextual cue, allowing an organism to adapt to different perturbations in diverse contexts. Here we show that trained singers' can rapidly acquire multiple internal models regarding voice fundamental frequency (F_0) that accommodate different perturbations to ongoing auditory feedback. Participants heard three musical notes and reproduced them in succession. Although participants were not explicitly instructed, the musical targets could serve as a contextual cue to indicate the direction feedback would be altered on each trial. Adaptation was observed immediately following vocal onset when participants were gradually exposed to altered feedback. Aftereffects were target specific and did not influence vocal productions on subsequent trials. When target notes were no longer a contextual cue, adaptation occurred during altered feedback trials and evidence for trial-by trial adaptation was found. These findings indicate that the brain is exceptionally sensitive to deviations between auditory feedback and the predicted consequence of a motor command during vocalization. Moreover, these results indicate that when contextually cued the vocal control system may

maintain multiple internal models that are capable of independent modification during different tasks or environments.

Introduction

Whether we are using a new tool, learning to speak another language, or singing a new song, the nervous system relies on various forms of feedback to establish task-specific sensorimotor representations. Over time, the plasticity of the nervous system permits neural reorganization and the formation of an 'internal model'. The prevailing hypothesis regarding the control of limb dynamics (Wolpert and Kawato 1998) and the control of speech (Guenther and Perkell 2004; Houde and Jordan 1998; Jones and Munhall 2005) and singing (Jones and Keough 2008; Keough and Jones 2009) is that internal models regulate motor movements. These internal representations are often investigated by altering a particular aspect of the sensory feedback associated with a given task. Results show that participants compensate by adjusting their movement in the opposite direction of the perturbation (Jones and Munhall 2000, 2005; Kalenscher et al. 2003; Sainburg et al. 1999; Shadmehr and Mussa-Ivaldi 1994). In adaptation studies, where dynamic manipulations are held fixed for a series of trials, aftereffects are often observed when feedback is unexpectedly returned to normal; subsequent responses err in the direction of compensation (Ghahramani and Wolpert 1997; Jones and Keough 2008; Keough and Jones, 2009; Shadmehr and Moussavi 2000). Furthermore, aftereffects have been observed while participants' perform untrained movements following exposure to altered feedback (Jones and Keough 2008; Jones and Munhall 2005; Keough and Jones 2009; Shadmehr and Moussavi 2000; Shadmehr and Mussa-Ivaldi 1994). This suggests that sensorimotor recalibration generalizes to other novel motor

productions. However, the arm reaching literature notes that the degree of generalization is reduced as the untrained movement direction diverges from the trained movement direction (Sainburg et al. 1999). Given that the transfer of aftereffects deteriorates proportionately as the distance from the trained location increases, it has been argued that precise movements correspond to context-specific neural mappings or multiple internal models (Kalenscher et al. 2003).

Recent evidence supports the existence of multiple internal models for the motor control of arm reaching movements (Donchin et al. 2003; Kalenscher et al. 2003; Osu et al. 2004; Wada et al. 2003; Wainscott et al. 2005; Wolpert and Kawato 1998). For example, individuals can acquire and appropriately switch between multiple internal representations for the same direction of movement when provided with a contextual cue (e.g., colour) (Osu et al. 2004; Wada et al. 2003; Wainscott et al. 2005). Currently, it remains unknown whether the vocal control system maintains multiple internal representations. The purpose of the present study was to investigate the sensitivity of trained singers' internal representations to altered auditory feedback and determine whether participants can develop and switch between multiple internal representations based on a contextual cue (musical notes). To identify whether voice F_0 is represented by multiple internal representations that are capable of independent recalibration and whether a contextual cue differentially affects the representations for F_0 control, trained singers ($n = 30$) produced three (Experiment 1, $n = 15$) different sequential target notes (A4, G4 and F4; 440, 392 and 349 Hz, respectively) or a

single (Experiment 2, $n = 15$) target note (G4), while receiving unaltered and frequency-altered feedback (FAF).

Materials and Methods

Participants. Experiments 1 and 2 had 30 Wilfrid Laurier University students (all women) whose native tongue was North American English participate in this frequency-altered feedback study. All participants were trained singers (mean musical training was approximately 12 years) recruited from the Faculty of Music at Wilfrid Laurier University. Participants received financial compensation for their contribution to this research. Informed consent was collected from each participant and the Wilfrid Laurier University Research Ethics Committee approved the procedures.

Apparatus. The experiment took place in a double-walled sound attenuated booth (Industrial Acoustic Company, Model 1601-01). Participants were fitted with a condenser microphone (Countryman Isomax E6 Omnidirectional Microphone), maintained approximately 3 cm from their mouth, and headphones (Sennheiser HD 280 Pro). In order to reduce natural acoustic feedback and bone-conducted feedback, participants heard multitalker babble noise (Auditec, St. Louis, MO) at 80 dB SPL while vocalizing. The target notes were produced by a trained female singer who sang the consonant-vowel /ta/. The target notes were processed using the speech modification algorithm STRAIGHT (Kawahara et al., 1999) to ensure that each target was exactly 349 (F4), 392 (G4) or 440 (A4) Hz. Microphone signals were directed to a signal processor (VoiceOne 2.0, TC Helicon) that manipulated auditory feedback. The

manipulated feedback was mixed (Mackie ONYX 1640) with the multitalker babble and presented to the participant. Vocal productions were digitized at 44.1 kHz for future analysis.

Procedure. Singers were instructed to emulate the musical target that was presented to them on each trial (total trials = 270). In Experiment 1 ($n = 15$) the first target was either A4 (440 Hz) or F4 (349 Hz), which was followed by the second target G4 (392 Hz) and then followed by the third target F4 or A4, respectively. Thus, the two conditions were A4-G4-F4 (135 trials) and F4-G4-A4 (135 trials), which was counterbalanced across singers. Participants received unaltered auditory feedback during the first 10 trials (baseline) for each target note in order to acclimatize them to the task and establish a baseline representation. Note that participants were presented with target notes in a sequential order, but had to reproduce them one at a time when prompted by a cue (the target note). For example, on the first trial they would be presented with the target A4, which they would reproduce to the best of their ability. On the second trial they would be presented with G4, and then on the third trial they would receive F4. This sequence would then be repeated (A4, followed by G4, then F4) until they produced the targets on all trials (135) per condition. Following baseline (training phase), auditory feedback for the first target note in the sequence (A4 or F4) was gradually manipulated (over 25 trials) upward in pitch (in 4 cent increments to 100 cents; 100 cents = 1 semitone). Thus, on the first trial with FAF for the initial target in the sequence, auditory feedback would be manipulated upwards by 4 cents. During the second FAF trial for the same note,

auditory feedback would be altered by 8 cents, and this would continue until auditory feedback was shifted by 100 cents (the last FAF trial for that target note). Auditory feedback for the second target (G4) in the sequence remained unaltered during the 25 training trials. G4 was used as a control note to examine the sensory motor representation of an unaltered target while compensating for FAF. Lastly, participants' auditory feedback for the third note in the sequence (F4 or A4) was gradually manipulated (over 25 trials) downward in pitch (in -4 cent increments to -100 cents). Thus, on the first trial with FAF for the third target in the sequence, auditory feedback would be manipulated downwards by -4 cents. During the second FAF trial for the same note, auditory feedback would be altered by -8 cents, and this would continue until auditory feedback was shifted by -100 cents (the last FAF trial for that target note). Although the targets were a contextual cue, participants were not informed of the relationship between the target note and the specific direction auditory feedback was manipulated during the training phase. During the final 10 trials (test phase) for each note participants received unaltered feedback that was used to examine F_0 productions following FAF.

The procedure for Experiment 2 ($n = 15$) was exactly the same as described above, however trained singers were only required to produce a single target (G4, 392 Hz) on all trials (total trials = 135). In doing so, the target notes were no longer a contextual cue however the direction of pitch shift manipulations remained predictable across FAF trials. Thus, the condition for this group was G4-G4-G4, where the first note in the sequence would be unaltered

for the first 10 trials (baseline), then gradually shifted upward in pitch over the 25 training trials (in 4 cent increments to 100 cents), followed by 10 trials (testing) with unaltered auditory feedback. The second note remained unaltered (across all baseline, training and test trials = 45), and the third note would be unaltered for the first 10 trials (baseline), then gradually shifted downward in pitch over the 25 training trials (in 4 cent increments to -100 cents), followed by 10 trials (testing) with unaltered auditory feedback.

An individual trial commenced with the presentation of a tone (1000 Hz) lasting 100 ms in duration, which warned participants that the trial was about to begin. Following the beep a 500 ms interstimulus interval preceded the target (A4, G4, or F4), which lasted 2000 millisecond. Only one target note was presented per trial and it occurred on all trials. Participants immediately received 4000 ms of multitalker babble after the presentation of the target and were instructed to emulate the note as accurately as possible in pitch and duration during this time. Thus, participants had to hold their vocal productions constant for approximately 2000 ms while trying to match the pitch of their voice to that of the target presented per trial. Note that participants' auditory feedback was presented at approximately 85 dB SPL whereas the multitalker babble was presented at approximately 75 dB SPL. F_0 values were calculated for each vocal production using an autocorrelation algorithm included in the Praat program (Boersma, 2001). F_0 values were normalized to each target note (F4, G4 or A4) by calculating the appropriate cent values using the following formula:

$$\text{Cents} = 100 (12 \log_2 F/B)$$

Where F is the F_0 value in Hertz and B is frequency of the target pitch participants were instructed to sing (349, 392 or 440 Hz).

Statistical Analyses. Only the initial 1500 ms of each vocal utterance was analysed, as previous research has identified compensatory responding during FAF typically occurs between 130 to 500 ms post perturbation onset (Burnett et al., 1997; Jones and Keough, 2008). The data for the singers F_0 values during the A4-G4-F4 and F4-G4-A4 sequences (Experiment 1) were analysed together and were broken down into blocks of five trials within each sequence: baseline trials (6-10), shift trials (1-5, 6-10, 11-15, 16-20, 21-25) and test trials (1-5, 6-10). An ANOVA was conducted on the mean F_0 values with 2 (sequence: A4-G4-F4 or F4-G4-A4) x 3 (pitch shift: upward, unaltered and downward) x 8 (block) as factors. The data for singers in the G4-G4-G4 condition (Experiment 2) were also broken down into blocks of five trials within each sequence: baseline trials (6-10), shift trials (1-5, 6-10, 11-15, 16-20, 21-25) and test trials (1-5, 6-10). An ANOVA was conducted on the mean F_0 values with 3 (pitch shift: upward, unaltered and downward) x 8 (block) as factors. Tukey's honestly significant difference (HSD) test was implemented for post-hoc analyses with an alpha level of .05 used for all statistical tests.

Typically, during FAF studies, researchers have only examined aftereffects following a series of altered feedback trials. However, if participants are altering an internal representation that permits them to compensate for these perturbations while receiving FAF, then examining the median value within the initial 50 ms of vocal onset should identify whether sensorimotor adaptation

occurred. The median F_0 values for each sequence were categorized into blocks of five trials in the same fashion as the mean F_0 values. Therefore, an ANOVA was carried out on the median F_0 values obtained in Experiment 1 during the initial 50 ms of vocalization with 2 (sequence: A4-G4-F4 or F4-G4-A4) x 3 (pitch shift: upward, unaltered and downward) x 8 (block) as factors. An ANOVA was also conducted on the median F_0 values obtained in Experiment 2 during the initial 50 ms of vocalization with 3 (pitch shift: upward, unaltered and downward) x 8 (block) as factors.

Results

Experiment 1 Singers' mean F_0 values were calculated for each condition and are depicted in Figure 1. The analysis of the mean F_0 values during the A4-G4-F4 and the F4-G4-A4 sequences resulted in a main effect of pitch shift ($F_{(2,28)} = 251.55, p < .05$). The mean F_0 values (A4 and F4) obtained during upward and downward pitch shift conditions were significantly lower and higher than the mean F_0 values obtained during the unaltered (G4) condition, respectively ($p < .05$). A significant pitch shift by block interaction ($F_{(14, 196)} = 167.61, p < .05$) revealed that singers' baseline F_0 values were significantly higher than F_0 values obtained during FAF blocks 2-5 during the upward shifted feedback trials (all p 's $< .05$). Also, singers' baseline F_0 values were significantly lower than the F_0 values observed during FAF blocks 2-5 of the downward shifted feedback trials (all p 's $< .05$).

Surprisingly, aftereffects were not found on trials following FAF for either upward or downward pitch shifted conditions (all p 's $> .05$). When auditory

feedback suddenly returned to normal during test trials singers adjusted their F_0 to levels not different than those obtained during baseline. Furthermore, baseline F_0 values for the unaltered target G4 were not significantly different than any other block of unaltered trials ($p > .05$), nor were they different than the baseline values of A4 and F4 during both the upward and downward pitch shift conditions (all p 's $> .05$). However the baseline F_0 values for G4 were significantly higher than F_0 values for FAF blocks 2-5 during upward pitch shift trials (all p 's $< .05$) and were significantly lower than downward pitch shift FAF blocks 2-5 (all p 's $< .05$).

Singers' median F_0 values during the first 50 ms of each utterance was calculated for each sequence and is presented in Figure 2. The analysis of the median F_0 values during the initial 50 ms of vocal onset during the A4-G4-F4 and F4-G4-A4 conditions produced a main effect of pitch shift, $F_{(2,28)} = 22.78$, $p < .05$. The median F_0 values during downward pitch shifts were significantly higher than the F_0 values during the no shift and upward pitch shift conditions, respectively ($p < .05$). Moreover, there was a significant pitch shift by block interaction ($F_{(14, 196)} = 13.33$, $p < .05$), which revealed that singers' baseline F_0 values for the target notes A4 and F4 during the shift up condition were significantly higher than the F_0 values obtained during FAF blocks 5, 6 and both blocks of unaltered (test) trials following FAF ($p < .05$). In addition, singers' baseline F_0 values obtained for the targets A4 and F4 during downward pitch shifts were significantly lower than F_0 values obtained during FAF blocks 5, 6 and the first block of test trials following FAF ($p < .05$). Thus, examining the initial 50 ms of vocal onset identified that

singers not only exhibited aftereffects when producing targets with upward and downward manipulations in auditory feedback, but that sensorimotor adaptation occurred online during blocks of FAF trials. Finally, generalization effects carried over to the F_0 values for the unaltered target following FAF trials. Baseline F_0 values for the unaltered target (G4) were significantly different than the first block of F_0 values following FAF ($p < .05$).

Experiment 2 Singers' mean and median (within 50 ms of vocal onset) F_0 values were calculated for the G4-G4-G4 condition and are depicted in Figure 3. Participants' mean F_0 values during the G4-G4-G4 condition yielded a main effect of pitch shift ($F_{(2,28)} = 414.95, p < .05$). Similarly to the results obtained in Experiment 1, trained singers' mean F_0 values collected during upward and downward pitch shift manipulations were significantly lower and higher than the mean F_0 values found for the unaltered target, respectively ($p < .05$). Also, a significant pitch shift by block interaction was observed ($F_{(14, 196)} = 171.04, p < .05$). Post-hoc testing revealed similar results as those in Experiment 1, that participants' baseline F_0 values were significantly higher than F_0 values obtained during all FAF blocks during the upward shifted feedback trials (all p 's $< .05$). Moreover, singers' baseline F_0 values were significantly lower than the F_0 values obtained during all FAF blocks of the downward shifted feedback trials (all p 's $< .05$).

Aftereffects were not observed during test blocks of trials when producing the target following FAF that was shifted either upward or downward in frequency (all p 's $< .05$). The mean F_0 values obtained from singers producing the target

with unaltered auditory feedback were similar across all blocks of trials ($p > .05$). Thus, the results of the singers' mean F_0 values are virtually identical to those obtained in Experiment 1. Participants compensated for FAF by altering their voice F_0 in the opposite direction of the manipulations, and producing a target(s) with FAF did not influence the F_0 values for an unaltered pitch target. Thus, providing a contextual cue results in a similar pattern of compensation whether participants produce three different target notes or the same target note with FAF.

Singers' median F_0 values during the initial 50 ms of vocal onset revealed a main effect of pitch shift, $F_{(2,28)} = 12.58$, $p < .05$. The median F_0 values during upward pitch shifts was significantly higher than the F_0 values obtained during the no shift and downward pitch shift conditions, respectively ($p < .05$). There were no differences found between the F_0 values in the no shift and downward pitch shift conditions ($p > .05$). The difference found between the F_0 values in the upward pitch shift condition compared to those in the unaltered and downward pitch conditions is opposite to that found in Experiment 1. It may be the case that producing the same target note with FAF and unaltered feedback resulted in trial-by-trial adaptation.

Indeed, post-hoc results of the significant pitch shift by block interaction ($F_{(14, 196)} = 13.33$, $p < .05$) revealed that trial-by-trial adaptation occurred for the unaltered pitch target. The F_0 values obtained during baseline were significantly different than the F_0 values obtained during the last three blocks of FAF trials ($p < .05$). Also, there was a trend, albeit nonsignificant and in the opposite direction of

the data obtained in Experiment 1 (See Figure 3), in the F_0 values during the blocks of FAF trials where auditory feedback was shifted downwards. Thus, it appears that the trial-by-trial adaptation observed on unaltered trials carried over to trials when auditory feedback was shifted downward, however the differences were not significant. Interestingly, the opposite pattern of aftereffects found for the main effect of pitch shift in Experiment 2 might be accounted for by trial-by-trial adaptation, given that trials with upward pitch manipulations immediately followed the trials with downward pitch manipulations. Overall, removing the potential to use target notes as a contextual cue resulted in similar compensatory responses, however the pattern of aftereffects observed appears to be the result of trial-by-trial adaptation.

Discussion

The data in Experiment 1 represents the first demonstration that vocal control may be represented by multiple internal representations, and that participants' acoustic-motor mappings are capable of independent sensorimotor recalibration. Participants' initial F_0 productions, within 50 ms of onset, were consistently influenced by the perturbed feedback experienced on previous FAF trials. For instance, in order to continually produce the target notes accurately while receiving FAF, trained singers had to progressively modify their vocal productions. Thus, as participants' auditory feedback was incrementally shifted (± 4 cents) trial-by-trial, we observed corresponding changes in open-loop control followed by rapid online correction for pitch deviations. That is, participants' adjusted their F_0 in the opposite direction of the perturbation once

the new discrepancy was detected. Over time, participants' compensatory responses resulted in the gradual recalibration of multiple internal representations. Interestingly, the aftereffects observed during training (FAF trials) did not generalize to vocal productions immediately following altered feedback trials. Rather the aftereffects were unique to the frequency of the target that was presented every third trial.

Even though the recalibration of internal representations was limited to the pitch-shifted targets during FAF trials, aftereffects were observed within 50 ms of vocal onset for the unaltered target following the training period (Figure 2). Transferred aftereffects (generalization) to an unaltered stimulus have been observed in previous FAF (Jones and Keough 2008; Jones and Munhall 2005) and arm reaching investigations (Ghahramani et al. 1996; Shadmehr and Mussa-Ivaldi 1994). In the current study and in other work (Jones and Keough 2008; Jones and Munhall 2005) pitch-shift manipulations were gradually presented during FAF. When feedback returned to normal, participants heard their F_0 for altered notes 1 semitone different than it was on the previous trial. Thus, the single-trial aftereffects observed in the median 50 ms F_0 data during the test trials for the unaltered pitch target might have been the result of the sudden and large changes in auditory feedback following training.

The results of Experiment 1 were analogous to some of the studies that investigated multiple internal models for the motor control of arm reaching movements. However, determining whether multiple internal models exist in the arm reaching literature has yielded inconsistent results. For instance, if the task

was dependent on colour cues (e.g., room light colour), if trials were presented randomly, if movements were both dynamic transformations or were dependent on the same state variable, or if the temporal interval between internal model acquisition was less than four hours, multiple internal models were not acquired (e.g., Brashers-Krug and Shadmehr 1996; Gandolfo et al. 1996; Karniel and Mussa-Ivaldi 2002; Krakauer et al. 1999; Tong et al. 2002). Karniel and Mussa-Ivaldi (2002) did not find evidence to suggest that participants could acquire and switch between internal models for two alternating viscous force fields, even after participants performed these movements in 4 sessions over 4 days. Rather they argued that a single internal model was formed and employed when presented with sequential perturbations.

On the other hand, several studies have found that participants could acquire internal models when prompted by a contextual cue (Osu et al. 2004; Wada et al. 2003; Wainscott et al. 2005). For example, Wada et al. (2003) found that participants could learn and switch between two internal models for opposing viscous force fields presented randomly cued only by colour. Moreover, Osu et al. (2004) found that providing visual cues prior to movement initiation allowed participants to predictively switch between acquired motor programs. In both cases, it was argued that multiple internal models were formed under diverse conditions, including single-joint movements to 4 or 8 target locations while receiving assistive/resistive or rotational forces to the limb. Moreover, the results from both groups (Osu et al. 2004; Wada et al. 2003) support the MOSAIC model (Kawato and Wolpert 1998; Wolpert and Kawato 1998), which suggests that

many controllers are selected and acquired while gated by a responsibility signal. The responsibility signal is thought to be determined by the accuracy of prediction of the paired predictors (forward models) as well as from contextual information. Therefore, the responsibility signal in MOSAIC decides the switching in addition to the discriminative learning of multiple internal models (for a review see Wada et al. 2003). However, the goodness of prediction made by several forward models alone is not fast enough to be utilized during online switching and for the selective learning of multiple inverse models while performing reaching movements in adaptation studies (Wada et al. 2003). This results from the fact that feedback is only available following the initiation of arm reaching movements. Thus, in order to compute the responsibility signal in time, Wada et al. (2003) argue that contextual information plays an essential role.

Our data supports the notion that contextual information is important in the acquisition and switching of multiple internal representations for vocal control while singing. We found that trained singers could rapidly acquire and independently modify multiple internal representations when cued by different target notes. However, unlike the aforementioned arm reaching studies, participants were not informed that the target notes could be used as a contextual cue. On the one hand, as long as the contextual information is clear and distinct (e.g., colour, target notes) then learning multiple environments can occur relatively easily (Wada et al. 2003). Indeed, the results of previous arm reaching studies (Osu et al. 2004; Wada et al. 2003) supporting multiple internal representations appear to be consistent with this interpretation. On the other

hand, if the contextual information is ambiguous or not present at all and if the multiple environments are difficult to discriminate (Brashers-Krug and Shadmehr 1996; Gandolfo et al. 1996; Karniel and Mussa-Ivaldi 2002; Krakauer et al. 1999; Tong et al. 2002), then acquiring or switching between multiple internal representations is difficult (Wada et al. 2003). Indeed, when the contextual cue was removed in Experiment 2, a completely different pattern of aftereffects was observed.

When the target notes could no longer be used as a contextual cue, to indicate the direction of pitch shift manipulation on the current FAF trial, we found trial-by-trial adaptation and similar patterns of compensatory responding as those found in Experiment 1. However it should be noted that the trial-by-trial adaptation was limited to the F_0 values corresponding to an unaltered pitch target. On trials following upward pitch shifted FAF, participants received natural acoustic feedback. Singers initiated vocal productions as if they were anticipating similar shifts in auditory feedback that were presented to them on the previous trial. When participants produced the same note in Experiment 1 with unaltered feedback no aftereffects were observed following FAF trials. Thus, it seems that the pattern of sensorimotor recalibration is dependent on the nature of the motor commands associated with the task.

In the case of arm reaching studies, participants have been required to initiate movements from a fixed location (Imamizu and Kawato 2008; Krakauer et al. 1999; Osu et al. 2004; Wada et al. 2003). This has permitted the examination of feed-forward internal models within 250 ms of movement initiation (e.g.,

Wainscott et al., 2005), which has been argued to be a period of time where motor commands (trajectories) are little influenced by closed-loop control or 'online' feedback. A unique aspect of studying voice F_0 while singing is that is not necessary (or actually possible) to require participants to initiate motor commands at a consistent starting point (a particular pitch). Moreover, previous data from our laboratory has also found that trained singers initiate vocal productions at or near the desired target frequencies while singing (Keough and Jones, 2009). Thus, we have been able to identify sensorimotor adaptation that occurs within 50 ms of vocal onset, and measure this adaptation over the course of training with dynamic perturbations cued by different target notes (see Figure 2).

Interestingly, previous FAF studies have relied exclusively on trials immediately following pitch-shifted feedback to examine sensorimotor adaptation (Jones and Keough 2008; Jones and Munhall 2000, 2005). Although our results confirmed that participants compensated for FAF, no aftereffects were observed in the mean F_0 data when feedback was returned to normal (see Figure 1). Conversely, on the same trials participants' initial F_0 values within 50 ms of vocal onset indicated that they expected the FAF, however, when no error was perceived they compensated by rapidly adjusting their F_0 to produce the target notes at frequencies not significantly different from baseline. This explains why there were no significant differences between singers mean F_0 values obtained during baseline and test trials. Thus, it may be the case that investigating the

effects as they occur online during altered feedback trials may provide a more sensitive measure of adaptation.

The compensatory responses observed were consistent with those obtained in previous FAF studies examining speech (Donath et al. 2002; Houde and Jordan 1998; Jones and Munhall 2000, 2005) and singing (Burnett et al. 1997; Jones and Keough 2008; Keough and Jones 2009; Natke et al. 2003; Toyomura et al. 2007; Zarate and Zatorre 2008). Regarding speech, previous studies have found that the speech motor system appears to be represented in a task-specific manner (Shaiman and Gracco 2002; Tremblay et al. 2008). For instance, Shaiman and Gracco (2002) found that applying unanticipated mechanical loads to the upper lip during speech production only influenced productions that required the upper lip (e.g., 'apa', 'p' requires both lips). Perturbing articulators uninvolved in the specific speech sounds being produced (e.g., 'afa', 'f' requires the lower lip) elicited no differences in electromyographic (EMG) activation between control and load trials (Shaiman and Gracco 2002).

Indeed, our data suggests that the motor system involved in the regulation of voice F_0 while singing is also organized in a task-specific manner. Aftereffects did not carry over to influence notes on subsequent trials following FAF, rather the aftereffects in Experiment 1 were limited to the notes participants produced every third trial. Tremblay and colleagues (2008) also argued that speech learning is contextually sensitive and generalization was not observed even when utterances shared similar movements. Although there is some degree of overlap when singers produced the target stimuli in our study, our data suggests that the

pitch of musical notes is not influenced by altered feedback experienced on previous trials and that singing may be represented by multiple frequency-specific internal representations.

Our work complements previous findings in the arm reaching literature studying multiple internal models for motor control. These data suggest that producing multiple target notes while singing requires participants to employ unique motor commands for each target. Although the human voice has the potential to be initiated at unpredictable frequencies during onset, trained singers consistently produced vocal pitch near the desired target frequencies, even in the presence of FAF. Interestingly, when the target notes no longer served as a contextual cue we observed a very different pattern of adaptation. Overall, our data suggests that sensorimotor adaptation is automatic, it can be observed during training within 50 ms of vocal onset while singing, and is uniquely associated with the motor commands for specific musical targets.

Acknowledgments

We thank M. Jarick for critical discussions and review of the manuscript. This research was supported by the National Institute of Deafness and Communicative Disorders Grant and a grant from the Natural Sciences and Engineering Research Council of Canada.

Figure Legends

Figure 1

Average single trial voice fundamental frequency (F_0) values across an entire utterance before, during and following frequency-altered feedback. Trials presented in red and blue represent the last baseline trial and the initial test trial following frequency-altered feedback (FAF) with unaltered auditory feedback, respectively. Trials presented in green represent the last FAF training trial with (a) auditory feedback manipulated upwards by 1 semitone (100 cents), (b) unaltered auditory feedback (control session), and (c) auditory feedback manipulated downwards by 1 semitone (-100 cents). The median F_0 value between 0 and 50 ms (vertical dotted line) was used to track sensorimotor adaptation. The mean F_0 values were calculated by averaging all data points across 1500 ms of vocal productions and were used to indicate compensatory responding in voice F_0 .

Figure 2

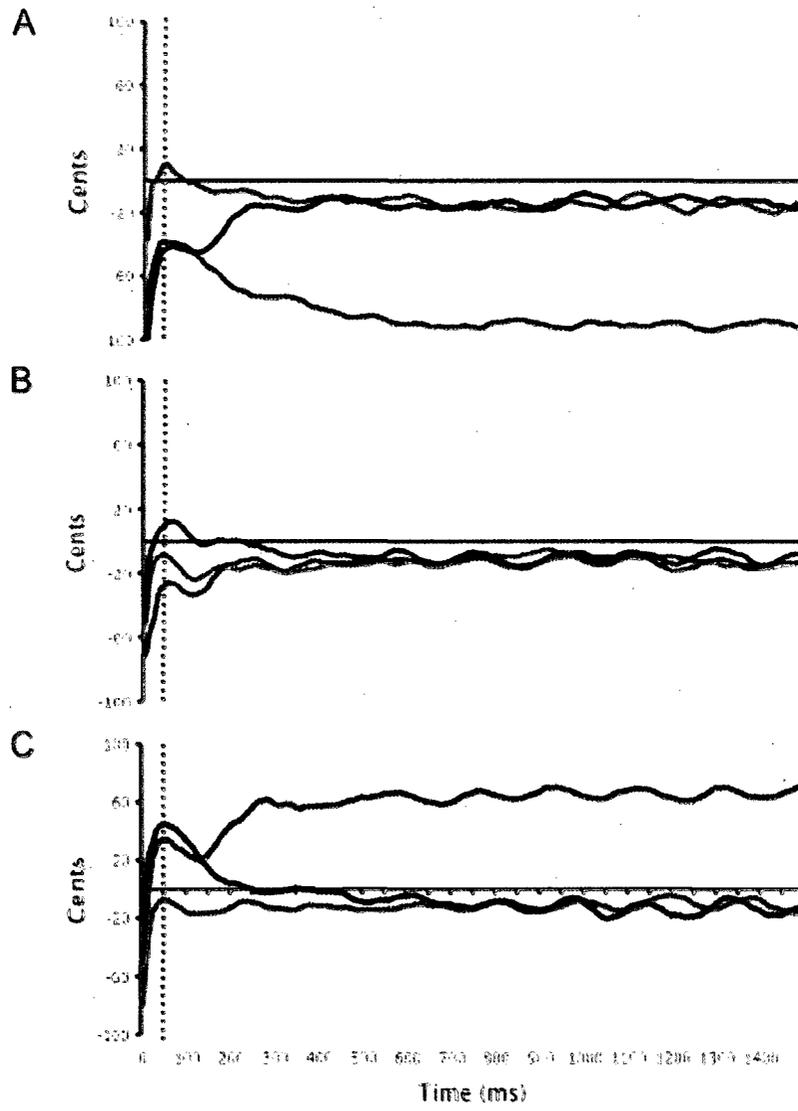
Average fundamental frequency (F_0) values obtained in Experiment 1 during blocks of frequency-altered feedback (FAF) and test trials. F_0 was calculated based on median value between 0 and 50 ms of vocal onset, or mean F_0 across 1500 ms of vocal productions. Data was normalized by subtracting the average of the last five baseline F_0 values from the F_0 values collected during FAF training and test trials. Vertical dotted line indicates the conclusion of FAF trials and the commencement of test trials (feedback returned to normal). Asterisks indicate significant post-hoc (Tukey's HSD) results for F_0 values when different than

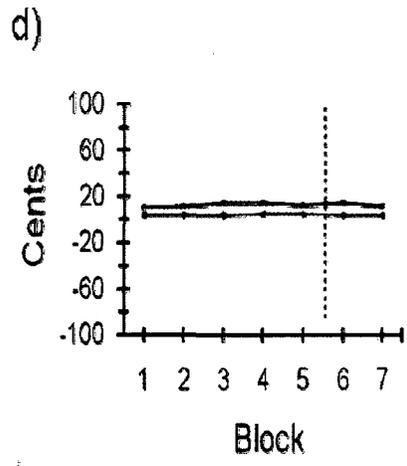
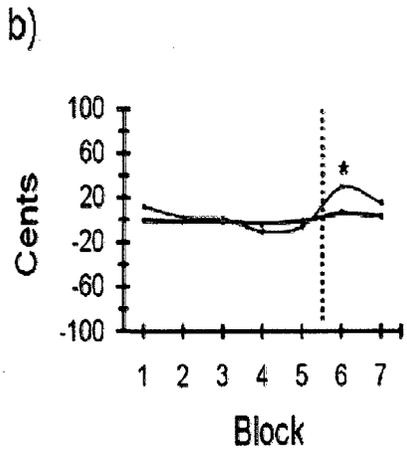
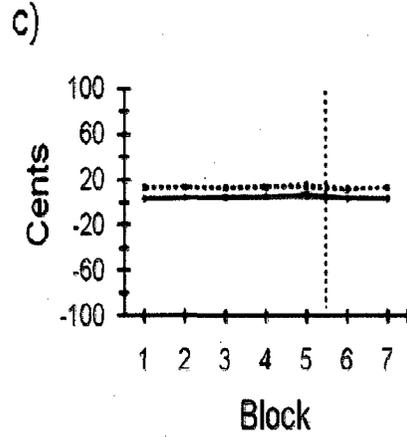
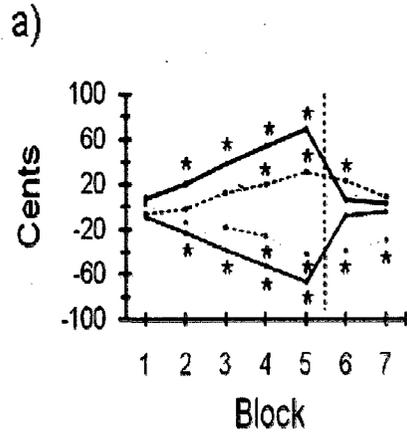
baseline. (a) Red and blue circles represent mean F_0 values and red and blue squares represent median 50 ms F_0 values when auditory feedback was unaltered. (b) Black circles represent mean F_0 values and grey squares represent median 50 ms F_0 values when auditory feedback was unaltered. (c) Standard error of the means presented in Figure 2a are presented. (d) Standard error of the means presented in Figure 2b are presented.

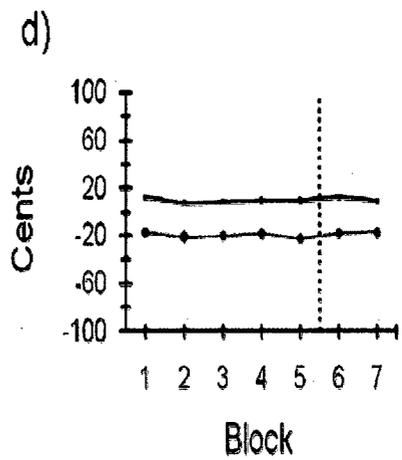
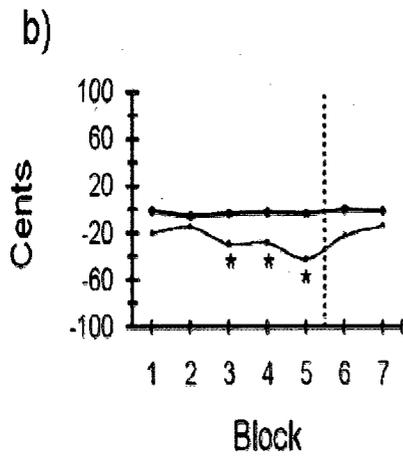
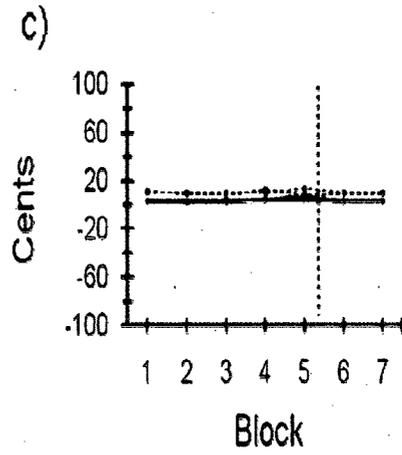
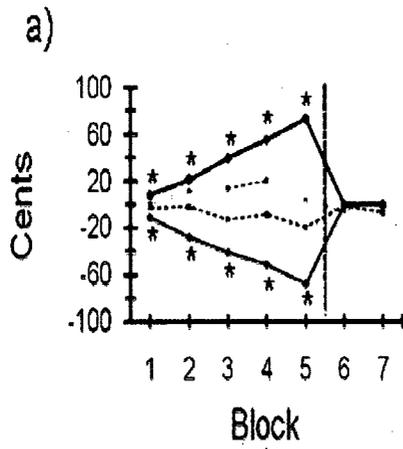
Figure 3

Average fundamental frequency (F_0) values obtained in Experiment 2 during blocks of frequency-altered feedback (FAF) and test trials. (a) F_0 was calculated based on mean F_0 across 1500 ms of vocal productions or (b) median value between 0 and 50 ms of vocal onset. Data was normalized by subtracting the average of the last five baseline F_0 values from the F_0 values collected during FAF training and test trials. Vertical dotted line indicates the conclusion of FAF trials and the commencement of test trials (feedback returned to normal).

Asterisks indicate significant post-hoc (Tukey's HSD) results for F_0 values when different than baseline. Black circles represent mean F_0 values and grey squares represent median 50 ms F_0 values when auditory feedback was unaltered. (c) Standard error of the means presented in Figure 3a are presented. (d) Standard error of the means presented in Figure 3b are presented.







Chapter 3

Running head: COMPENSATION AND SENSORIMOTOR ADAPTATION

**Evidence of Compensation and Sensorimotor Adaptation even when
Participants were Instructed to Ignore Frequency-Altered Feedback**

Dwayne Keough¹, Colin Hawco¹, & Jeffery A. Jones ^{*1,2}

¹ Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

² Centre for Cognitive Neuroscience
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

*Author to whom correspondence should be addressed:

Jeffery A. Jones
Department of Psychology
Wilfrid Laurier University
Waterloo, Ontario
N2L 3C5
Canada

E-mail: jjones@wlu.ca
Tel: 519-884-0710 ext: 2992
Fax: 519-746-7605

Abstract

Auditory feedback has been shown to be influential in the maintenance and control of voice fundamental frequency (F_0). The purpose of this study was to address whether task instructions could influence the compensatory responding and sensorimotor adaptation that has been previously found when participants are presented with a series of frequency-altered feedback (FAF) trials. Trained singers and musically untrained participants (nonsingers) were informed that their auditory feedback would be manipulated in pitch while they sang the target vowel [/ah/] (as in 'pop'). Participants were instructed to either 'compensate' or 'ignore' the changes in auditory feedback. Whole utterance auditory feedback manipulations were either gradually presented ('ramp') in -2 cent increments down to -100 cents (1 semitone) or were suddenly ('constant') shifted down by 1 semitone. Results indicated that singers and nonsingers could not suppress the compensatory responses to FAF, nor could they reduce the sensorimotor adaptation observed on both ramp and constant FAF trials. Compared to previous research these data suggest that musical training is effective in suppressing compensatory responses only when FAF occurs after vocal onset (500-2500 ms). Moreover, our data suggest that compensation and adaptation is automatic and is influenced little by conscious control.

195 Words

Keywords: compensate, ignore, internal model, frequency-altered feedback, auditory feedback, fundamental frequency, pitch, musical training, singing, voice, singer.

1. Introduction

The goal of both speech and singing is the production of sound. With respect to the acoustical properties of both, the fundamental frequency (F_0), or vocal pitch, may be one of the most salient features. For instance, violating the target pitch contour of the vocal production may affect the meaning of the utterance (Grell et al., 2009). Arguably the most influential contribution to F_0 control is from auditory feedback. Several studies have shown that auditory feedback is involved in the control of voice F_0 (Burnett et al., 1997, 1998; Elman, 1981; Houde & Jordan, 1998; Jones & Keough, 2008; Jones & Munhall, 2000, 2002; Keough & Jones, 2009; Lane & Tranel, 1971; Munhall et al., 2009). For instance, modifying auditory feedback generally results in compensatory responses in participants' ongoing vocal productions.

There is little doubt that sensory feedback is essential for the acquisition and maintenance of precise motor control (e.g., Burnett et al., 1997, 1998; Elman, 1981; Ghahramani & Wolpert, 1997; Houde & Jordan, 1998; Kalenscher et al., 2003; Keough & Jones, 2009; Jones & Keough, 2008; Jones & Munhall, 2000, 2002; Larson et al., 2008; Munhall et al., 2009; Sainburg et al., 1999; Shadmehr & Moussavi, 2000; Shadmehr & Mussa-Ivaldi, 1994). However, recent evidence (Finney & Palmer, 2003; Zarate & Zatorre, 2008) questions the influence of various forms of sensory feedback, specifically auditory feedback, on ongoing motor productions when participants possess extensive training with the task. For instance, Finney and Palmer (2003) demonstrated that presenting or removing auditory feedback did not differentially affect the quality of pianists'

performance when asked to execute a well-rehearsed piece from memory. Similarly, Zarate and Zatorre (2008) found that trained singers could almost completely ignore auditory feedback perturbations (± 200 cents, 2 semitones) and continually produce ('sing') the intended target similarly to when auditory feedback was unaltered. On the other hand, musically untrained participants were unable to suppress the compensatory responses to the frequency-altered feedback (FAF) and adjusted their vocal productions in the opposite direction of the perturbations. Likewise, masking auditory feedback with noise led to a greater number of errors singing pitch targets in nonsingers than in trained singers (Schultz-Coulton, 1978; Ward & Burns, 1978). Thus it seems that musical training may contribute to musicians' precise performance, even when auditory feedback is altered, masked, or eliminated altogether. One possibility is that the differences observed in the control of voice F_0 were accomplished by an increased reliance on an 'internal model'.

The feedback associated with particular movements are thought to be represented by an 'internal model'. Internal models are proposed to exist as neural maps of skilled movements that store the relationships between the motor commands, environment and sensory feedback for their production (Desmurget & Grafton, 2000; Flanagan & Wing, 1993; Shadmehr & Mussa-Ivaldi, 1994). In fact, the prevailing hypothesis regarding the control of limb dynamics (Wolpert & Kawato, 1998) and the control of speech (Houde & Jordan, 1998; Guenther & Perkell, 2004; Jones & Munhall, 2005) and singing (Jones & Keough, 2008; Keough & Jones, 2009) is that internal models regulate motor movements. Given

the considerable training singers possess, it was thought that an internal model for vocal control would be more entrenched in singers than in untrained participants (nonsingers). Jones and Keough (2008) had trained singers and nonsingers reproduce a musical target (392 Hz or 349 Hz, G4 or F4 respectively) while receiving unaltered or FAF that was shifted down by one semitone (100 cents). Although participants compensated for the FAF, aftereffects were only observed in singers' data. Singers' F_0 values during testing were higher than their baseline F_0 values, whereas nonsingers' F_0 values during testing were similar to those obtained during baseline (Jones & Keough, 2008). Singers' data also indicated that aftereffects generalized to other notes than the one they sang during training (FAF trials). Indeed, one possible explanation for the difference in voice F_0 control is that singers' internal models could recruit similar neural substrates as nonsingers but in varying degrees, or alternatively they are recruiting different neural substrates altogether.

Recently, Zarate and Zatorre (2008) used functional magnetic resonance imaging (fMRI) to assess the neural processes associated with singing under different feedback conditions. When participants were presented with normal auditory feedback and were instructed to sing various target notes as accurately as possible, the behavioural results indicated that singers were more accurate and less variable in producing the musical targets than nonsingers (Zarate & Zatorre, 2008). Despite the differences in vocal production during the normal feedback condition, both singers and nonsingers exhibited similar functional networks for singing. These areas included the bilateral auditory cortices,

bilateral primary motor cortices, the supplementary motor area (SMA), the anterior cingulate cortex (ACC), thalamus, insula, and cerebellum (Zarate & Zatorre, 2008). Unlike the compensatory responses reported by Jones and Keough (2008), Zarate and Zatorre did not find behavioural differences when instructing participants to compensate for the FAF. It should be noted however that participants were not informed of the pitch shift modifications in Jones and Keough (2008), whereas participants in Zarate and Zatorre (2008) were informed that their auditory feedback would be altered in pitch. On the other hand, Zarate & Zatorre (2008) did find enhanced activation in the ACC, STS, insula, putamen, pre-SMA, and IPL in singers. Lastly, when Zarate and Zatorre exposed participants to FAF and instructed them to ignore their feedback (i.e., do not compensate for the manipulations), the fMRI results revealed that singers showed enhanced activation in the inferior parietal lobule (IPL), superior temporal gyrus (STG), superior temporal sulcus (STS), and right insula (Zarate & Zatorre, 2008). Interestingly, the pattern of behavioural results obtained suggests that singers can suppress compensatory responses to FAF, providing the shifts occur between 1000-1500 ms following vocal onset. Essentially, singers' F_0 values during the ignore condition were similar to those obtained when they received unaltered auditory feedback. On the other hand, nonsingers were unable to ignore the FAF and compensated by altering their F_0 in the opposite direction of the perturbations.

The results of Zarate and Zatorre (2008) raise an interesting question about the influence of musical training and the use of auditory feedback on voice

F_0 control; is it possible that participants' compensatory responding is under conscious control? According to the results previously mentioned, it seems that as long as participants possess sufficient vocal training then they may have been able to consciously suppress compensatory responses to FAF by relying on alternative feedback mechanisms (e.g., muscle memory for pitch, proprioception) during these trials. However, when Munhall and colleagues (2009) presented participants with formant frequency manipulations that coincided with vocal onset, a robust compensation was observed in all conditions even when they instructed speakers to ignore changes in auditory feedback. Moreover, although a number of studies conducted by Larson and colleagues (Burnett et al., 1997, 1998; Burnett & Larson, 2002; Larson et al., 2001; but see Hain et al., 2000) were not directly investigating whether participants could ignore FAF, their results suggest that participants compensate for pitch shift manipulations, even when told to keep their voice stable and to ignore any auditory feedback variations presented over the headphones. These data suggest that compensation to formant perturbations (Munhall et al., 2009) and FAF (Burnett et al., 1997, 1998; Burnett & Larson, 2002; Larson et al., 2001) is automatic and that compensatory responses do not appear to be modified by a conscious strategy. The only difference previously reported was that F_0 response latencies were reduced when participants had immediate control over perturbation onset (Burnett, McCurdy, & Bright, 2008). No differences were observed to the direction, magnitude, or the peak time of voice F_0 responses (Burnett et al., 2008).

Indeed, how conscious awareness influences the perception and production of speech and singing is unknown. The results of Loui et al. (2008) suggest a distinction between pathways for vocal performance and conscious perception may exist. Thus, possessing knowledge of the forthcoming perturbations does not appear to alter the direction, magnitude, or peak time of F_0 responses. This suggests that the pathway involved with the conscious detection of altered feedback may not be required for eliciting compensatory responses. Even though the majority of participants in previous studies were not musically trained, or the extent of their training was not well established, it could be that the extensive training singers receive may provide them with the ability to utilize alternative forms of feedback more efficiently to suppress compensatory responses to FAF.

The current study will examine singers and nonsingers F_0 responses to gradual (-2 cent increments to -100 cents, 1 semitone) or constant (-100 cents) changes in FAF. We will be particularly interested in whether participants can suppress compensatory responses to FAF manipulations when instructed to do so. That is, we will examine participants' vocal responses when instructed to compensate or maintain (ignore FAF) a steady voice F_0 during FAF trials to determine if compensatory responses can be influenced by volitional control. Moreover, we will also investigate whether aftereffects (adaptation) can be suppressed during FAF trials. Aftereffects have been typically observed following a series of FAF trials (Jones & Keough, 2008; Jones & Munhall, 2000, 2002), where vocal productions err in the direction of compensation. However recent

evidence from our laboratory identified that sensorimotor adaptation can be observed within 50 ms of vocal onset (Hawco & Jones, 2009; Keough & Jones, 2009). Surprisingly, during a recent adaptation study in our laboratory (Keough & Jones, 2009) we did not find aftereffects in singers' mean F_0 values (across 1500 ms of vocal production) following FAF, whereas we observed robust aftereffects in a previous study (Jones & Keough, 2008). As a result, we decided to examine participants F_0 values within 50 ms of vocal onset. During trials following FAF, we observed adaptation followed by rapid compensation, such that singers initiated their F_0 values as if they were anticipating the FAF and when it was determined that their auditory feedback had been returned to normal, they adjusted their mean F_0 to values consistent with those obtained during baseline (prior to FAF). Therefore, singers in Zarate and Zatorre's (2008) study may have exhibited sensorimotor adaptation, but rapidly altered their F_0 to suggest they could ignore the manipulations. As a consequence, the extent to which compensation and adaptation is under volitional control remains unknown.

The current study had singers and nonsingers produce target notes at specific frequencies with different instructions while receiving subtle and large modifications in auditory feedback. Participants were informed that their auditory feedback was manipulated in pitch and they were instructed to either (1) 'compensate' for these changes by altering their voice F_0 in the opposite direction of the perturbation or (2) to 'ignore' their auditory feedback maintain their voice F_0 at frequencies similar to when their feedback was unaltered. The purpose is to investigate (i) whether task instructions influence compensatory responding (F_0

values within 1500 ms of vocal onset) and sensorimotor adaptation (F_0 values within 50 ms of vocal onset) that are typically observed during frequency-altered feedback studies (e.g., Burnett et al., 1997, 1998; Jones & Keough, 2008; Jones & Munhall, 2000, 2002; Keough & Jones, 2009). Moreover, (ii) collecting data from singers and nonsingers will help identify whether musical training influences acoustic-motor control when instructed to compensate or ignore auditory feedback manipulations. Regardless of whether the pitch manipulations are small (-2 cent increments down to -100 cents, where 100 cents = 1 semitone) or large (-100 cents), if both singers and nonsingers can suppress or eliminate compensatory responses and sensorimotor adaptation then it will suggest that these responses are, to a certain degree, under volitional control. Conversely, if similar patterns of compensatory responding and sensorimotor adaptation are observed, then it will suggest that these responses are automatic and that auditory feedback has an important role in vocal motor control while singing.

It is hypothesized that both singers and nonsingers will be unable to ignore subtle shifts (-2 cent increments to -100 cents; 'ramp condition') in FAF. As a result, it is expected that participants will exhibit similar patterns of compensatory responding and sensorimotor adaptation during the ramp condition. When auditory feedback is suddenly shifted to -100 cents ('constant condition'), it is hypothesized that nonsingers will immediately compensate by increasing their voice F_0 in the opposite direction of the manipulation. Over time, it is believed that nonsingers will exhibit sensorimotor adaptation while compensating for the FAF. On the other hand, trained singers may be able to

ignore large changes in auditory feedback when instructed to do so, which would be congruent with the results of Zarate and Zatorre (2008). Indeed, if singers can ignore FAF manipulations then their vocal productions should be similar to those obtained when auditory feedback is unaltered (baseline). However a key methodological difference is that the perturbations coincided with vocal onset in this study, whereas Zarate and Zatorre (2008) presented FAF between 1000-1500 ms following vocal onset. This is particularly important because Hawco and Jones (2009) found that presenting FAF manipulations prior to vocalization elicited larger compensatory responses than those observed when FAF occurs shortly after vocal onset. It is possible that the suppression of compensatory responding exhibited by trained singers in Zarate and Zatorre (2008) resulted from their reliance on the auditory feedback they received prior to the presentation of the FAF. Thus, presenting participants with subtle and large changes in auditory feedback that coincides with vocal onset may indicate that compensatory responding and sensorimotor adaptation occurs automatically and is influenced little by volitional control.

2. Methods

2.1 Participants

Thirty Wilfrid Laurier University students (all women) whose native tongue was North American English participated in this frequency-altered feedback study. Currently, no evidence suggests that a gender difference exists in relation to vocal control while receiving FAF. Men were not included so that all

participants could adequately sing the same target notes. Of the 30 participants, 15 were trained singers recruited from the Faculty of Music (vocal majors) at Wilfrid Laurier University (mean music training was approximately 12 years). None of the trained singers reported having 'perfect' pitch. The remaining 15 participants were nonsingers, such that none had any form of previous vocal training or were currently participating in formal singing. All participants received financial compensation for their time and informed consent was collected from each participant. The Wilfrid Laurier University Research Ethics Committee approved the procedures.

2.2 Apparatus

Participant Recording Sessions. Participants were situated in a double-walled sound attenuated booth (Industrial Acoustic Company, Model 1601-01) and were fitted with headphones (Sennheiser HD 280 Pro) and a condenser microphone (Countryman Isomax E6 Omnidirectional Microphone), which was positioned approximately 3 cm from their mouth. Multitalker babble noise (Auditec, St. Louis, MO) was presented at 80 dB SPL (sound pressure level) to limit natural acoustic feedback. Note that the multitalker babble was unintelligible to the listener, as it consisted of 20 young adults simultaneously reading different passages. The target notes consisted of a female voice singing the vowel-consonant [/ah/], that was presented at 220.00 (A3), 246.94 (B3), 293.66 (D4) or 329.63 (E4) Hz, respectively. Microphone signals were sent to a signal processor (VoiceOne 2.0, TC Helicon) that manipulated auditory feedback. The altered

feedback was then mixed (Mackie ONYX 1640) with the multi-talker babble and subsequently sent to the participant. Vocal productions were digitally recorded (TASCAM HD-P2) at a sampling rate of 44.1 kHz for future analysis.

Target stimuli recording. A trained singer produced the respective targets, A3, B3, D4, and E4, which were processed using the speech modification algorithm STRAIGHT (Speech Transformation and Representation using the Adaptive Interpolation of weighted spectrum; Kawahara et al., 1999) to ensure that each target was exactly 220.00, 246.94, 293.66, or 329.63 Hz.

2.3 Procedure

Participants were asked to match the pitch of their voice to a musical target during 4 conditions that consisted of 320 trials, which were divided into 4 blocks of 80 trials. Each block consisted of 30 baseline and 50 FAF trials (see Figure 1 for an outline of the methods). For instance, on one block participants would reproduce the musical target A3 (220.00 Hz) on all trials. Thus, participants would receive natural acoustic (unaltered) feedback during the 30 baseline trials, followed by 50 FAF trials. During the FAF trials, auditory feedback was either gradually shifted downward (ramp condition) in -2 cent increments per trial down to -100 cents, or it was simply shifted down (constant condition) 100 cents for all 50 FAF trials. Note that auditory feedback was shifted from the beginning of each utterance until the end of vocal productions during FAF trials. Another block of trials had participants reproduce B3, D4 or E4 for 30 baseline and 50 FAF trials. Target notes were presented in the following order for all

participants, A3-B3-D4-E4. However whether participants received FAF that was ramped (incrementally by 2 cents) or constantly (-100 cents) manipulated was counterbalanced across participants. Moreover, whether participants received instructions to compensate or ignore the FAF was also counterbalanced across participants.

Participants were instructed to either 'compensate' or 'ignore' any changes in auditory feedback that may occur during the study. During the 'compensate' condition, participants were informed that they would be presented with a musical target at the beginning of each trial. The goal for participants was to match the pitch of their voice as accurately as possible to the target presented. Participants were informed that they would initially receive normal, unaltered, auditory feedback. Such that what they produced would be exactly what they would hear in the headphones. Additionally, they were told that at some point the pitch of their voice presented via the headphones would be different than what they actually produced. Thus, they were instructed to continually monitor their auditory feedback and to try to match what they heard in the headphones to the target presented at the beginning of each trial. Essentially, participants were instructed to compensate for the FAF.

During the 'ignore' condition, participants were informed that they would be presented with a musical target at the beginning of each trial. The goal for participants was to match the pitch of their voice to the target presented, but to ignore their auditory feedback. In other words, they were told that initially they would receive normal, unaltered, auditory feedback. However, they were

informed that at some point the pitch of their voice presented via the headphones would be different than what they actually produced. Participants were instructed to ignore the FAF (their auditory feedback) and to produce the target consistently, at the same pitch as when their feedback was unaltered. Essentially, we were asking participants to 'not compensate' for the pitch shift manipulations, so sounding 'off' would be acceptable during these trials. At no time during the study were the participants informed of the direction of the FAF manipulation, at what trial the perturbation occurred, or made aware of alternative strategies that could be used to assist with vocal control (e.g., vocal fold position).

An individual trial commenced with the presentation of the target stimulus, which lasted for 1000 ms. Immediately following the termination of the target, multitalker babble was presented for 3500 ms. During the presentation of the multitalker babble, participants were instructed to reproduce the target as accurately as possible in pitch and duration. That is, hold the pitch of their voice constant for approximately 1000 ms. A 1000 Hz beep coincided with the last 500 ms of the multitalker babble, which served to signify that the trial was about to conclude and participants should cease vocalization. An intertrial-interval (ITI) was presented for 3000 ms between trials. Trials were initiated and controlled by a computer and F_0 values for each vocal production were calculated, during offline analyses, using an autocorrelation algorithm included in the Praat program (Boersma, 2001). F_0 values were normalized to each target note (A3, B3, D4 or E4) by calculating the appropriate cent values using the following formula:

Cents = $100 (12 \log_2 F/B)$

Where F is the F_0 value in Hertz and B is frequency of the target pitch

participants were instructed to sing (220.00, 246.94, 293.66, or 329.63 Hz).

2.4 Statistical Analyses

The median value within the initial 1500 ms of each vocal utterance was analysed, as it has been previously shown that compensatory responses to FAF typically occurs between 130 to 500 ms post perturbation onset (Burnett et al., 1997). Participants F_0 values obtained during the 'ignore' and 'compensate' conditions were analysed together and were broken down into blocks of five trials: baseline trials (1-5, 6-10, 11-15, 16-20, 21-25, 26-30), and FAF trials (1-5, 6-10, 11-15, 16-20, 21-25, 26-30, 31-35, 36-40, 41-45, 46-50). Thus, an ANOVA was conducted on the median F_0 values with 2 (group: singers and nonsingers) x 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) as factors. Tukey's honestly significant difference (HSD) test was used for post-hoc analyses with an alpha level of .05 for all statistical tests.

Moreover, unlike previous adaptation studies (e.g., Jones & Keough, 2008; Jones & Munhall, 2000, 2002, 2005) examining aftereffects following a series of FAF trials, we recently identified that sensorimotor adaptation can be observed within 50 ms of vocal onset (Hawco & Jones, 2009; Keough & Jones, 2009). As a result, the median F_0 values within 50 ms of vocal onset for each sequence were categorized into blocks of five trials in the same fashion as the median 1500 ms F_0 values. Thus, an ANOVA was conducted on the median 50

ms F_0 values with 2 (group: singers and nonsingers) x 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) as factors.

Singers and nonsingers 50 ms data was also subjected to a similar ANOVA however the group variable was no longer included in the analysis. Therefore, for both singers and nonsingers a 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) ANOVA was conducted to determine the degree of sensorimotor adaptation that occurred during FAF.

3. Results

The median 1500 ms F_0 values for singers and nonsingers were calculated for each condition and are displayed in Figure 2. The ANOVA conducted on the median 1500 ms F_0 values with 2 (group: singers and nonsingers) x 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) as factors revealed a marginal effect of group, $F(1, 28) = 3.17, p = .086$. On average, singers' median 1500 ms F_0 values were lower, or more consistently near the desired target frequency, than nonsingers' median 1500 ms F_0 values. A significant main effect of manipulation and block was also observed, $F(1, 28) = 28.08, p < .05$ and $F(15, 420) = 210.64, p < .05$, respectively. Overall, the F_0 values obtained for all participants during the ramp manipulation were significantly lower than the F_0 values obtained during the constant manipulation. Thus, participants compensated to a greater degree when presented with a constant manipulation (-100 cents) than when presented with gradual shifts in auditory feedback (-2 cent increments to -100 cents). With

respect to the main effect of block, post-hoc testing indicated that all baseline F_0 values collected during both ramp and constant manipulation conditions were not significantly different (all p 's > .05). However, the average F_0 values obtained during baseline trials for both ramp and constant manipulations were found to be significantly less than the average F_0 values obtained during all FAF blocks of trials for both ramp and constant manipulation conditions, (all p 's < .05).

A significant two-way interaction between manipulation and block was also found, $F(15,420) = 75.71, p < .05$. Post hoc analysis revealed that the average baseline blocks (1-6) of F_0 values obtained during the ramp manipulation condition were significantly different than the average F_0 values obtained on FAF blocks 9-16. As the pitch shift manipulation became progressively lower, participants, on average, compensated by increasing the pitch of their voice in the opposite direction of the manipulations so that the F_0 values were different than those during baseline. Also, as the pitch shift manipulations during the ramp condition progressively decreased, participants' F_0 values gradually became significantly different than those observed on previous blocks of FAF trials ($p < .05$). For instance, the F_0 values obtained on the last block (block 16) of FAF trials during the ramp manipulation condition were significantly higher than all F_0 values obtained on previous blocks of FAF trials ($p < .05$). With respect to the F_0 values obtained during the constant manipulation condition, post hoc testing indicated that the baseline values were significantly lower than the F_0 values obtained during all FAF blocks (7-16) of trials ($p < .05$). Interestingly, the F_0 values observed during the initial block (7) FAF trials was not statistically different

than the F_0 values observed on any other block (8-16) of FAF trials ($p > .05$). Therefore, singers and nonsingers, on average, compensated immediately and consistently to the sudden and constant change (-100 cents) in auditory feedback across all blocks of FAF trials. Although, the F_0 values collected on the initial block (7) during the constant manipulation condition were significantly higher than those values collected on FAF blocks 7-14 of the ramp manipulation condition (all p 's $< .05$), they were not significantly different than the F_0 values obtained on blocks 15 and 16 of the ramp condition. Thus it appears that the level of compensation to FAF is similar when auditory feedback is manipulated downward between 80-100 cents. Lastly, no significant effect of instruction (ignore or compensate) was observed in the median 1500 ms F_0 values, $F(1, 28) = .011$, $p < .05$. Even when instructed to ignore the FAF, participants were unable to maintain their voice F_0 at similar levels to those obtained during baseline. No other significant main effects or interactions were observed.

The median 50 ms F_0 values for singers and nonsingers were calculated for each condition and are displayed in Figure 3. The ANOVA conducted on the median 50 ms F_0 values with 2 (group: singers and nonsingers) x 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) as factors revealed a significant main effect of manipulation and block, $F(1, 28) = 8.80$, $p < .05$ and $F(15, 420) = 15.36$, $p < .05$, respectively. On average, participants' median 50 ms F_0 values during the ramp manipulation condition were significantly lower than the 50 ms F_0 values during the constant manipulation condition. Post hoc results on the main effect of block revealed that

participants F_0 values during baseline trials were not significantly different ($p > .05$). On the other hand, the F_0 values obtained during baseline (blocks 1-6) were found to be significantly lower than the F_0 values obtained during FAF blocks 12-16 (all p 's $< .05$). Moreover, some of the baseline F_0 values (blocks 3, 4 & 6) were significantly lower than the F_0 values observed on FAF blocks 8-16 (all p 's $< .05$). Thus, sensorimotor adaptation was observed when participants were subjected to FAF, regardless of whether they were instructed to compensate or ignore the manipulated auditory feedback.

A significant two-way interaction was also observed between manipulation and block, $F(15, 420) = 3.62, p < .05$. Participants' baseline F_0 values during the ramp manipulation condition were not statistically different nor were they different than the baseline F_0 values obtained during the constant manipulation condition (all p 's $> .05$). However, the F_0 values observed during baseline in the ramp condition were significantly lower than those values observed during FAF blocks 13, 15 and 16 ($p > .05$). Thus, as participants' auditory feedback was gradually shifted downward, they progressively increased their voice F_0 so that they initiated vocal productions at levels closer to the intended target frequency. Similar results were also observed in the median 50 ms F_0 values during the constant manipulation condition. Baseline F_0 values for blocks 2-6 were found to be significantly lower than the F_0 values on FAF blocks 8-16 (all p 's $< .05$). Interestingly, when auditory feedback is held fixed at a large pitch shift value (-100 cents) sensorimotor adaptation occurs rapidly and voice F_0 is maintained at consistent levels across FAF trials. Lastly, the main effect of instruction was not

significant, $F(1, 28) = .03$, $p > .05$. Participants' voice F_0 values were similar across all blocks of trials, regardless of whether they were instructed to compensate or ignore the FAF. No other significant main effects or interactions were observed.

In an attempt to further examine sensorimotor adaptation, independent ANOVA's were conducted on singers and nonsingers median 50 ms F_0 values with 2 (instruction: ignore or compensate) x 2 (manipulation: ramp or constant) x 16 (block) as factors. The analysis on nonsingers median 50 ms F_0 values revealed a significant main effect of block, $F(15, 210) = 4.42$, $p < .05$. Post hoc tests indicated that baseline blocks 1, 2, 5, and 6 were not significantly different than the F_0 values observed on FAF blocks 7-16 (all p 's $> .05$). However, baseline F_0 values on blocks 3 and 4 were significantly different than the F_0 values on FAF blocks 9, 12, 13, 15 and 16 (all p 's $< .05$). It appears that, on average, while sensorimotor adaptation is evident in nonsingers, they tend to consistently initiate voice F_0 values at a similar frequency. No other significant main effects or interactions were observed.

Results of singers median 50 ms F_0 values revealed a main effect of manipulation, $F(1, 14) = 13.56$, $p < .05$. The F_0 values obtained during the ramp manipulation condition were significantly lower than the F_0 values obtained during the constant manipulation condition. A significant main effect of block was also found, $F(15, 210) = 22.09$, $p < .05$. On average, post hoc testing indicated that all baseline F_0 values were significantly lower than the F_0 values on FAF blocks 10-16 (all p 's $< .05$). Unlike nonsingers, trained singers appear to consistently adjust their initial vocal productions in the presence of FAF, such

that their voice F_0 is initiated more closely to the target frequency. Thus, sensorimotor adaptation could be observed, on average, across most blocks of FAF trials during the ramp and constant manipulation conditions. Interestingly, sensorimotor adaptation could be observed earlier on baseline blocks 2, 5 and 6, as they were found to be significantly different than FAF blocks 8-16. Also, a significant manipulation x block interaction was observed, $F(15, 210) = 4.76$, $p < .05$. Post hoc testing during the ramp manipulation condition revealed that singers' baseline median 50 ms F_0 values for blocks 4-6 were significantly lower than the F_0 values found on FAF blocks 13-16. Moreover, baseline F_0 values for block 2 were significantly lower than FAF blocks 13, 15 and 16. Baseline blocks 1 and 3 were only significantly lower than FAF block 15. Thus, evidence for sensorimotor adaptation exists, suggesting singers are particularly sensitive to subtle changes in auditory feedback and compensate by altering how they initiate subsequent vocal productions. On the other hand, post hoc results on the constant manipulation condition revealed that singers' baseline F_0 values during all baseline blocks were significantly lower than the F_0 values on FAF blocks 9-16. Additionally, the level of sensorimotor adaptation observed within 50 ms of vocal onset during the last 2 FAF blocks of the ramp manipulation condition was not significantly different than any FAF block (7-16) during the constant manipulation condition. Thus, the degree of sensorimotor adaptation for F_0 values when auditory feedback was being gradually shifted between 80-100 cents was similar to the sensorimotor adaptation observed during the constant

manipulation condition. No other significant main effects or interactions were found.

4. Discussion

The purpose of the present study was to determine whether instructing participants to 'compensate' or 'ignore' gradual (-2 cent increments per trial down to -100 cents) or constant changes (-100 cents) in auditory feedback could result in the voluntary suppression of compensatory responding and sensorimotor adaptation. Regardless of whether participants received the gradual or constant pitch manipulations, both singers and nonsingers could not intentionally suppress the compensatory motor commands during FAF trials (see Figure 3). The pattern of compensation observed when participants were instructed to 'ignore' the FAF was indistinguishable from the compensatory responses observed when they were instructed to 'compensate' for the FAF. Additionally, participants' median 50 ms F_0 values suggested that the level of sensorimotor adaptation that occurred during the ignore condition was similar to the adaptation observed during the compensation condition. Voice F_0 values observed throughout FAF (gradual and constant) trials indicates that both singers and nonsingers updated their internal models by adjusting their F_0 so that they initiated vocal productions at frequencies closer to the intended target.

However, when participants' median 50 ms data was analysed separately, it was found that nonsingers were more likely to initiate voice F_0 at consistent frequencies. That is, nonsingers F_0 values during most blocks of FAF trials were

similar to those obtained when auditory feedback was unaltered. This does not suggest that nonsingers were able to suppress the sensorimotor adaptation typically observed during FAF trials. Rather this finding is consistent with previous data from our laboratory suggesting that nonsingers' internal models for vocal control while singing are not as well established as those observed in trained singers (Jones & Keough, 2008; Keough & Jones, 2009). For instance, in Keough and Jones (2009) it was found that nonsingers initiated voice F_0 approximately a semitone below the target frequency and rapidly corrected (increased) the pitch of their voice, during both unaltered and altered feedback trials. On the other hand, singers' F_0 values suggested that they continually updated their internal models in the presence of FAF to initiate vocal pitch near the desired target frequency (within 50 ms of singing). Indeed, trained singers median 50 ms data in the current study is consistent with this finding. Thus, trained singers tend to adjust their internal models in the presence of FAF to allow them to sing the target note immediately (within 50 ms of vocal onset), whereas nonsingers tend to employ a searching strategy to match their voice to the target. Note that nonsingers also modify an internal model to accommodate their searching strategy, however the recalibration observed in nonsingers is not as apparent as it has been in singers F_0 responses (Keough & Jones, 2009). Essentially, nonsingers modify their starting point while singing with FAF but continue to increase the pitch of their voice to similar degrees as when their auditory feedback was unaltered.

On the other hand, when participants were instructed to 'compensate' for the gradual presentation of FAF they correspondingly adjusted their F_0 in the opposite direction of the manipulation. That is, within 1500 ms of vocal onset, both singers and nonsingers increased their voice F_0 in order to maintain pitch accuracy with the intended target (see Figure 2). Although nonsingers F_0 values were consistently higher than singers during baseline and in the presence of FAF, the differences were only marginally significant. Moreover, participants also exhibited sensorimotor adaptation while compensating for the gradual FAF manipulations. Participants' data within 50 ms of vocal onset indicates that as the pitch manipulations progressively decreased, they initiated vocal productions at increasingly higher frequencies. In other words, participants started subsequent utterances at F_0 values similar to those obtained on previous FAF trials. This suggests that as singers and nonsingers were gradually compensating for the FAF, they were also continually updating their internal models to account for the consistently decreasing changes in auditory feedback.

Similarly, when participants were instructed to compensate for the sudden and large change (constant condition) in auditory feedback, results indicated that participants appropriately increased their F_0 in the opposite direction of the manipulation. Interestingly, the F_0 values, on average, obtained during the first block of FAF trials were not statistically different than the F_0 values obtained on any other block of FAF trials. Thus, participants compensated for the large shift in auditory feedback at consistent values from the initial presentation to the conclusion of the 50 FAF trials. Furthermore, participants' median 50 ms F_0 data

identified that sensorimotor adaptation occurred. As singers and nonsingers rapidly compensated for the FAF, they also adjusted their internal models to initiate F_0 values closer to the intended target. F_0 values were determined to be significantly different than baseline from the second block (trials 6-10) of FAF trials onward. As a consequence, instructing participants to compensate for FAF results in similar responses to those observed previously in our laboratory (Keough & Jones, 2009) and by others using the FAF paradigm (Burnett et al., 1997, 1998; Donath et al., 2002; Jones & Munhall, 2000; 2002; Natke et al., 2003; Zarate & Zatorre, 2008).

The finding that compensatory responding is not easily suppressed by instructions to ignore feedback is consistent with previous studies using FAF (Hain et al., 2000), formant frequency manipulations (Munhall et al., 2009), and masking noise (Pick et al., 1989). Indeed, a recent study by Munhall and colleagues (2009), found that participants rapidly compensated for formant frequency manipulations when they were instructed to ignore the modified feedback. Moreover, when the manipulations were removed participants exhibited aftereffects. As a consequence, Munhall and colleagues (2009) argue that their data do not necessarily provide “evidence of a fixed-response system that cannot be adjusted with practice or strategies” (pp. 389), but rather argue that compensatory responses to vowel modifications are not intentional strategic responses to the detection of auditory feedback manipulations. This is also congruent with the findings from the current study, however it is uncertain whether repeated exposure (‘practice’) to subtle (2 cents) manipulations in

auditory feedback would result in the overt suppression of compensatory responses to FAF.

For instance, when similar pitch shift values were presented incrementally across trials in previous studies from our laboratory (± 2 cents and ± 4 cents) participants stated that they were unaware that their voice was manipulated in pitch. Munhall and colleagues also indicated that participants possessed no particular knowledge of the nature of the manipulation when formants were modified in small increments trial-by-trial (e.g., Percell & Munhall, 2006). Indeed, it has been reported that an early automatic response to unexpected changes in auditory feedback occurs between 100-150 ms post perturbation onset (Burnett et al., 1997, 1998; Hain et al., 2000). Thus, if this response assists with small, unexpected, perturbations then the presentation of gradual shifts in auditory feedback may fall within a certain automatic compensatory range that cannot be overtly suppressed, nor may it require the 'conscious' detection of the error for the compensatory response to occur. This is consistent with the results of Loui et al. (2008), who reported that amusic ('tone-deaf') participants were able to reproduce the pitch direction of two successive single tones, although they were at chance discriminating pitch direction. That is, although amusics have difficulty perceptually identifying pitch changes that are smaller than a semitone (Peretz & Hyde, 2003), they are capable of producing the correct pitch direction as accurate as controls (Loui et al., 2008). This supports Loui et al.'s (2008) notion that the auditory pathway responsible for vocal production may be distinct from the pathway responsible for conscious perception. Thus, compensating for

altered feedback may occur without participants' ability to consciously identify this behaviour. Alternatively, repeated exposure to large changes (e.g., 100, 200 cents) in auditory feedback may allow compensatory responses to be overtly controlled (e.g., Hain et al., 2000; Zarate & Zatorre, 2008). Regardless, the data presented by Munhall et al. (2009) and the results of the current study suggest that motor preparation, initiation, and production of vocal utterances are heavily influenced by auditory feedback. Moreover, instructing participants to ignore changes in feedback does not appear to influence compensatory responding or alter the pattern of sensorimotor adaptation.

Auditory feedback has been shown to be important for accurate F_0 control, and it has also been shown to be influential during the acquisition of a novel musical piece. Finney and Palmer (2003) found that trained pianists performance was improved when auditory feedback was provided while learning a novel song. However, when the musicians were required to produce a well-rehearsed piece from memory, the removal of auditory feedback had no affect on performance (Finney & Palmer, 2003). Similar to the trained singers in Zarate and Zatorre (2008) who could suppress compensatory responses to +/-200 cents (2 semitone) manipulations, it appears that musical training may contribute to musicians' ability to perform in the absence or modification of auditory feedback. In regards to singing, one possibility is that presenting the pitch manipulations so they occur later into vocal production (1000-1500 ms in Zarate & Zatorre, 2008) may result in the easier identification of FAF (e.g., efference copy violation), or it may allow for singers to rely on alternative components (e.g., muscle memory,

kinesthetic feedback) of their internal model to suppress compensatory responding.

Conceptually, internal models are hypothesized to compare sensory feedback with motor acts by means of a comparator examining differences between perception and production. These differences are hypothesized to be computed based on a corollary discharge, such that the output of an internal model maps the motor commands (e.g., efference copy) with the expected sensory feedback from the actions. When a match exists between perception and production the result is a net cancellation of the sensory input, which in turn causes a dampened sensory experience (Heinks-Maldonado et al., 2005). Conversely, when there is a discrepancy between the perception and production of a motor act there is no corollary discharge to cancel the sensory feedback. As a consequence, there is an intensification of the sensory experience that potentially alerts us to environmental events (Heinks-Maldonado et al., 2005).

For instance, in a series of event related potential (ERP) and magnetoencephalographic (MEG) studies using FAF, Heinks-Maldonado and colleagues (2005, 2006) found that an early sensory detection component (e.g., M100) generated in the auditory cortex was maximally suppressed to participant's own unaltered voice. When participants received pitch-shifted and alien feedback conditions the researchers observed an increase in cortical responding in the auditory cortex relative to when they received unaltered auditory feedback (Heinks-Maldonado et al., 2006). According to this theory, participants in the current study should have also exhibited similar cortical

responding when presented with FAF, as they initiated compensatory responses. However, presenting the pitch manipulations so they coincide with vocal onset may make compensatory responding more difficult to suppress than if the FAF was presented during mid utterance. For example, when FAF is delivered mid utterance (e.g., Hain et al., 2000; Zarate & Zatorre, 2008), the efference copy associated with the motor commands is not initially violated, as the participant first hears exactly what they are producing. When the FAF occurs, it is possible that the nervous system has already determined that the motor commands are appropriate for the target note produced and that the error perceived is due to something external (e.g., the experimenter). Given the extensive experience trained singers possess with vocal control, participants in Zarate and Zatorre's (2008) study may have altered their internal model to rely more on kinesthetic feedback (e.g., vocal-fold positioning) to maintain the pitch of their voice during FAF trials, whereas nonsingers, possibly due to their lack of formal music training, were unable to suppress compensatory responses.

On the other hand, when the FAF coincides with vocal onset, the actual sensory feedback associated with the motor commands of participants' internal models does not match the expected sensory feedback, resulting in an intensified sensory experience (efference copy violation) (e.g., Heinks-Maldonado et al., 2005, 2006). Regardless of whether participants were 'consciously' aware of the FAF manipulations, compensatory responses were initiated to subtle and large changes in auditory feedback. This suggests that once the efference copy has been violated, participants' internal models are

automatically recalibrated and compensatory responses are initiated in an attempt to offset the auditory feedback manipulations.

Indeed, nonreflexive components (e.g., kinesthetic feedback, auditory feedback) are influential in achieving precise vocal control. Murbe et al. (2004) and Larson et al. (2008) have also demonstrated that kinesthesia substantially contributes to singers pitch control, but very early in responses (< 100 ms). After 100 ms auditory feedback participates in F_0 control (Larson et al., 2008). However, Munhall and colleagues (2009) recently found that instructing participants to rely on the kinesthetic properties for F_0 control was insufficient to suppress compensatory responding to formant frequency manipulations. If the participants in Munhall et al. (2009) were not musically trained then it may be that they were unable to utilize the kinesthetic feedback as efficiently as trained singers to suppress compensatory responses.

Indeed, our results and those of others (Hain et al., 2000; Jones & Keough, 2008; Keough & Jones, 2009; Munhall et al., 2009; Zarate & Zatorre, 2008) suggest that the processes involved in comparing the actual sensory consequences with the expected sensory consequences during vocalization is mainly dependent on auditory feedback. Singers' ability to ignore FAF (e.g., Zarate & Zatorre, 2008) may result from relying less on auditory feedback and more on supplementary feedback properties (e.g., kinesthetic) once a musical piece has been memorized (Finney & Palmer, 2003). Alternatively, a more likely explanation is that participants may utilize the information they receive following vocal onset differently to maintain a stable voice F_0 as opposed to the information

they receive mid utterance (Hawco & Jones, 2009). Overall, it appears that vocal training may only be effective in suppressing compensatory responses to FAF in instances when the perturbations are presented mid utterance.

Conclusion

Results from the present study suggest that neither musically trained singers, or nonsingers could overtly suppress compensatory responses to gradual or large FAF manipulations. Sensorimotor adaptation was also observed during both ignore and compensate conditions, however singers' data suggested an increased reliance on an internal model for vocal control. Formal music training appears to be useful in suppressing compensatory responses only when the FAF is presented following vocal onset (e.g., 500-2500 ms) (Hain et al., 2000; Zarate & Zatorre, 2008). Whether participants could suppress sensorimotor adaptation during mid utterance FAF manipulations remains unknown. In sum, it appears that compensation and adaptation to FAF are automatic and are influenced little by 'conscious' control, providing the FAF manipulations coincide with vocal onset.

Acknowledgments

The National Institute of Deafness and Communicative Disorders Grant DC-08092 and a grant from the Natural Sciences and Engineering Research Council of Canada supported this research.

Figure Caption

Figure 1

Auditory feedback shifts for two phases in the experimental conditions (gradual or constant frequency-altered feedback; FAF). During baseline (trials 1-30), auditory feedback was not manipulated in pitch. During perturbation (FAF) trials, presented in gray, auditory feedback was manipulated from trial 31 (solid vertical line) to trial 80 (all FAF trials represented by the dashed line). **(a)** Auditory feedback was gradually manipulated downwards, in -2 cent increments across trials until auditory feedback was shifted by -100 cents (1 semitone). **(b)** Auditory feedback was shifted by a constant value (-100 cents) for all FAF trials.

Figure 2

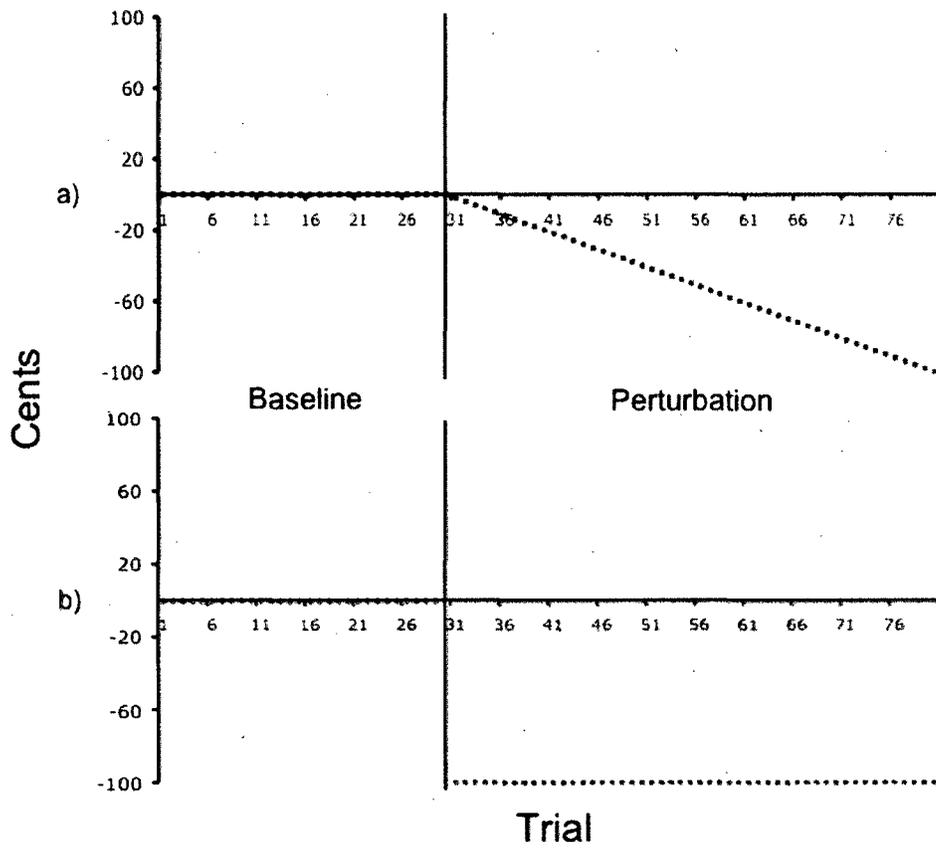
Average fundamental frequency (F_0) values across blocks (5 trials per block) of unaltered (baseline) and frequency-altered feedback (FAF) trials when participants were instructed to 'compensate' or 'ignore' changes in auditory feedback. F_0 was calculated based on median value across the initial 1500 ms of vocal productions. Vertical dotted line indicates the commencement of FAF trials. **(a)** Nonsingers median 1500 ms data across blocks of unaltered and FAF trials, where auditory feedback was gradually (-2 cent increments across trials to -100 cents) shifted in pitch. Black squares represent nonsingers F_0 values when instructed to 'compensate' for FAF. Gray diamonds represent nonsingers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented. **(b)** Singers median 1500 ms data across blocks of unaltered and FAF trials, where auditory feedback was gradually (-2 cent increments across trials to

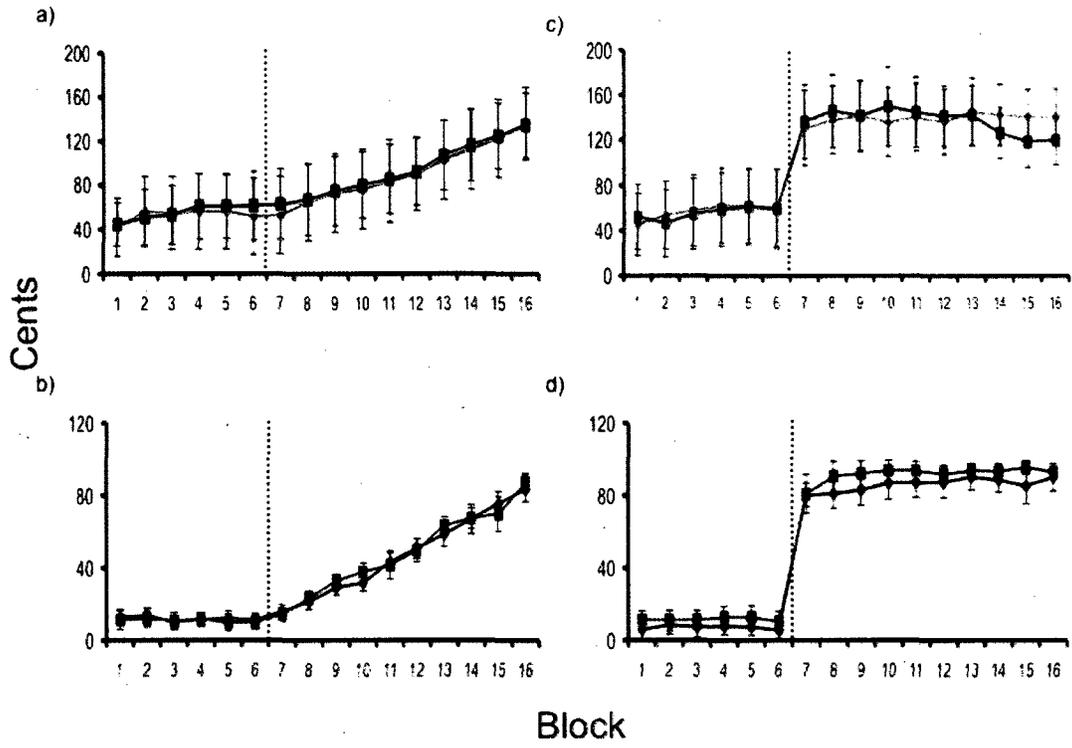
-100 cents) shifted in pitch. Red squares represent singers F_0 values when instructed to 'compensate' for FAF. Blue diamonds represent singers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented. (c) Nonsingers median 1500 ms data across blocks of unaltered and FAF trials, where auditory feedback was constantly shifted down by -100 cents. Black squares represent nonsingers F_0 values when instructed to 'compensate' for FAF. Gray diamonds represent nonsingers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented. (d) Singers median 1500 ms data across blocks of unaltered and FAF trials, where auditory feedback was constantly shifted down by -100 cents. Red squares represent singers F_0 values when instructed to 'compensate' for FAF. Blue diamonds represent singers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented.

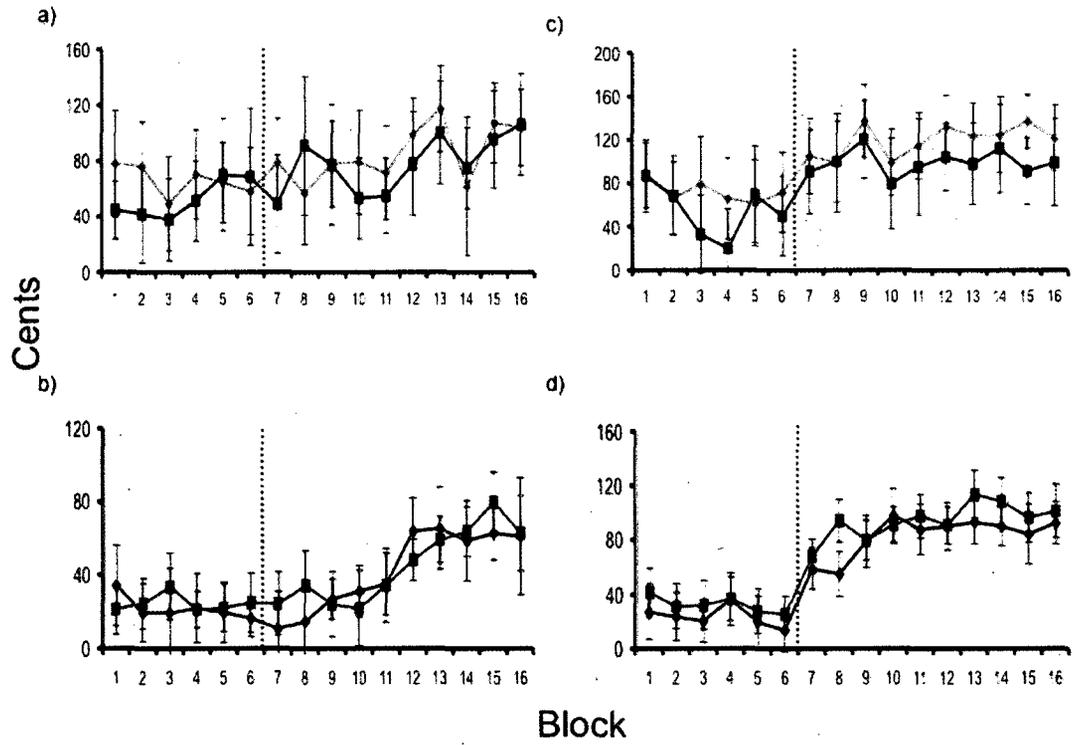
Figure 3

Average fundamental frequency (F_0) values across blocks (5 trials per block) of unaltered (baseline) and frequency-altered feedback (FAF) trials when participants were instructed to 'compensate' or 'ignore' changes in auditory feedback. F_0 was calculated based on median value across the initial 50 ms of vocal productions. Vertical dotted line indicates the commencement of FAF trials. (a) Nonsingers median 50 ms data across blocks of unaltered and FAF trials, where auditory feedback was gradually (-2 cent increments across trials to -100 cents) shifted in pitch. Black squares represent nonsingers F_0 values when instructed to 'compensate' for FAF. Gray diamonds represent nonsingers F_0

values when instructed to 'ignore' the FAF. Standard error of the means are presented. (b) Singers median 50 ms data across blocks of unaltered and FAF trials, where auditory feedback was gradually (-2 cent increments across trials to -100 cents) shifted in pitch. Red squares represent singers F_0 values when instructed to 'compensate' for FAF. Blue diamonds represent singers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented. (c) Nonsingers median 50 ms data across blocks of unaltered and FAF trials, where auditory feedback was constantly shifted down by -100 cents. Black squares represent nonsingers F_0 values when instructed to 'compensate' for FAF. Gray diamonds represent nonsingers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented. (d) Singers median 50 ms data across blocks of unaltered and FAF trials, where auditory feedback was constantly shifted down by -100 cents. Red squares represent singers F_0 values when instructed to 'compensate' for FAF. Blue diamonds represent singers F_0 values when instructed to 'ignore' the FAF. Standard error of the means are presented.







General Discussion

The goal of this dissertation was to determine whether auditory feedback is important for the control of voice fundamental frequency (F_0) while singing. Moreover, this research also addressed whether sensorimotor representations (internal models) regulate voice F_0 while singing. The answer to both these questions is undeniably yes. In order to adequately address the role of musical aptitude on the use of auditory feedback for vocal control, we had musically trained singers and nonsingers serve as participants. Using the frequency-altered feedback (FAF) paradigm we presented singers and nonsingers with manipulated auditory feedback and observed compensatory responses and sensorimotor adaptation on vocal productions.

Initially, I was interested in establishing whether participants would exhibit a similar pattern of responding to those observed in previous FAF studies examining speech production (Burnett et al., 1997, 1998; Burnett & Larson, 2002; Elman, 1981; Larson et al., 2001; Hain et al., 2000; Jones & Keough, 2008; Jones & Munhall, 2000, 2002, 2005; Zarate & Zatorre, 2005, 2008). However one goal of this dissertation was to extend the results to vocal control while singing. That is, if auditory feedback is important for the development and maintenance of the accurate control of voice F_0 while speaking, is it also the case that auditory feedback is crucial for the control of voice F_0 while singing?

Using Frequency-Altered Feedback to Examine F_0 Control While Singing

The majority of previous altered feedback studies examining speech production have found that participants will compensate for FAF by adjusting their F_0 in the opposite direction of the manipulation (Burnett & Larson, 2002; Burnett et al., 1997, 1998; Jones & Munhall, 2000, 2002, 2005; Liu et al., 2007; Natke et al., 2003). This is also true regardless of whether the pitch manipulations coincided with vocal onset (e.g., Heinks-Maldonado et al., 2005; Jones & Keough, 2008, Jones & Munhall, 2000, 2002, 2005) or were presented during mid utterance (e.g., Burnett et al., 1997, 1998; Hawco & Jones, 2009; Larson et al., 2001; Zarate & Zatorre, 2008). Our initial study (Jones & Keough, 2008) confirmed that both singers and nonsingers, like participants in speech production studies, compensated for FAF (-100 cents, 1 semitone) by modifying the pitch of their voice in the opposite direction of the manipulation. However, of particular interest was the degree to which participants compensated for the FAF when presented with an absolute target (a musical note), compared to when they are presented with a relative target, which was the case for a majority of studies examining speech production.

When singers and nonsingers were presented with a semitone (100 cents) decrease in auditory feedback we observed near perfect levels of compensation (Jones & Keough, 2008). This is congruent with the results obtained by Burnett et al. (1997), who reported exemplary data for one trained singer who exhibited near perfect levels of compensation to FAF while a singing musical scale. Moreover, Natke and colleagues (2003) found that participants, who were not trained singers, compensated more during the singing condition (66 cents) than

during the speaking condition (47 cents). Conversely, these data are in contrast to those observed during speech production, where participants have not exhibited complete levels of compensation during F_0 manipulations (e.g., Burnett et al., 1997; Donath et al., 2002; Larson, 1998; Larson et al., 2000; Natke et al., 2003). Rather, the response magnitudes reported have ranged from approximately 15-65 cents with an average response magnitude of around half a semitone, or 50 cents. It should be noted that a main methodological difference between studies of speech production and singing are that the former typically presented random perturbations (compensation study) to participants while they produced a target without reference to a specific frequency. The latter had participants receive a series of FAF trials (adaptation study) while trying to sing musical notes. Thus, having participants produce a series of targets that have a specific frequency may contribute to the more complete levels of compensation observed.

Aside from the compensatory magnitude differences, the results found in Jones and Keough (2008) also indicated that sensorimotor adaptation occurred on trials immediately following the removal of altered auditory feedback. Sensorimotor adaptation (aftereffects) has generally been observed following a series of perturbation trials when feedback suddenly and unexpectedly returns to normal (Ghahramani & Wolpert, 1997; Jones & Keough, 2008; Jones & Munhall, 2000, 2005; Shadmehr & Moussavi, 2000; Shadmehr & Mussa-Ivaldi, 1994). Interestingly though, sensorimotor adaptation was only observed in the data obtained from trained singers. Singers' F_0 values during testing (on trials

following FAF) were higher than their baseline F_0 values, whereas there were no differences observed in nonsingers' baseline and test data (Jones & Keough, 2008). Additionally, aftereffects generalized to another target note that singers did not produce during training (with FAF). Consistent with speech, these data were taken to assume that an internal model regulates voice F_0 while singing, which appears to be more entrenched in singers than in nonsingers.

The Sensitivity of Internal Models to Subtle Changes in Auditory Feedback

The prevailing hypothesis for the control of limb dynamics (Wolpert & Kawato, 1998), and the control of speech (Houde & Jordan, 1998; Guenther & Perkell, 2004; Jones & Munhall, 2000, 2005) and singing (Jones & Keough, 2008) is that internal models regulate motor movements. As mentioned in the general introduction, internal models are thought to exist as neural maps of skilled movement that store the relationship between the motor commands, environment and sensory feedback responsible for their production (Desmurget & Grafton, 2000; Flanagan & Wing, 1993; Shadmehr & Mussa-Ivaldi, 1994). Given that the results of Jones and Keough (2008) suggested an internal model corresponds to vocal control while singing, Chapter 1 was designed to assess the sensitivity of the mechanisms that modify internal representations in singers and nonsingers when presented with subtle changes in auditory feedback. Thus, we were interested in whether small, incremental changes in auditory feedback would have an effect on the control of voice F_0 control while singing.

The results presented in Chapter 1 demonstrate that singers and nonsingers compensated (mean 1500 ms data) to a similar degree when presented with subtle pitch shift manipulations. This is not surprising given that Watts and colleagues (2003) found that untrained singers were equally as accurate at producing pitch targets as trained singers. However, one difference between the groups in Chapter 1 was that singers are immediately and consistently more accurate when they reproduced the target notes. Nonsingers data suggested that they, on average, employed a 'searching' technique, by consistently initiating vocal productions below the target frequency ('flat') and rapidly increased the pitch of their voice to match the note as accurately as possible. This strategy was observed whether they were presented with auditory feedback that was manipulated upwards or downwards.

With respect to when participants initiated compensatory responses to FAF, singers' F_0 values suggested this occurred when feedback was manipulated by +/- 6 cents (1.36 Hz) during both the shift up and down conditions. This value is consistent with Sundberg's (1987) finding that trained singers can correct for production errors with an accuracy of less than 1 Hz from an intended pitch target (A4, 440 Hz). On the other hand, nonsingers initiated compensatory responses at approximately 24 cents (5.47 Hz) during shifted up and down conditions. These data are consistent with the just-noticeable difference range found by Pape and Mooshammer (2006) of F_0 contours for digitally synthetic stimuli. Moreover, nonsingers' values were very similar to the average threshold reported by Hafke (2008), who found that pitch shift changes

were not reliably identified when they were below 26 cents. However, Loui and colleagues (2008) found that control participants psychophysical thresholds of perception and production were around 2.0 and 2.5 Hz, respectively. Overall, this suggests that trained singers internal models are more sensitive to subtle changes in auditory feedback and they are capable of initiating compensatory responses to overcome very small errors in vocal production.

Sensorimotor Adaptation is Observable Within 50 ms of Vocal Onset

As outlined in the general introduction and in the results of Jones and Keough (2008), sensorimotor adaptation has been typically observed following a series of FAF trials, where vocal productions err in the direction of compensation. However, the results of Chapter 1 did not indicate that aftereffects occurred in singers' mean F_0 values (across 1500 ms of vocal production) following FAF. Indeed, the lack of aftereffects was quite unexpected. Essentially, we were uncertain whether adaptation had occurred at all during the study. For instance, it may have been the case that participants simply initiated vocal productions at similar frequencies and compensated online for the FAF. As a result, we decided to examine participants F_0 values within 50 ms of vocal onset in order to determine how participants initiated F_0 responses while singing with FAF. It was thought that music training may have resulted in an increased ability for singers to establish vocal fold position prior to producing sound than nonsingers (Watts et al., 2003). In fact, previous studies have found that trained singers are more accurate at prephonatory tuning, or positioning the laryngeal structures prior to

vocal onset, than nonsingers (Leonard & Ringel, 1979; Murry, 1990). Thus, trained singers should produce a steady and presumably a more accurate F_0 immediately from vocal onset than nonsingers. Moreover, if their internal models were recalibrated as the FAF gradually became higher or lower then we may correspondingly observe changes in successive vocal productions or, in other words, sensorimotor adaptation. Indeed, no study to date had found evidence to suggest the recalibration of an internal model that was observable within 50 ms of motor production.

Indeed, results from Chapter 1 showed that singers incorporate the discrepancy between perception and production by starting subsequent utterances during FAF trials where they ended on the previous production. Therefore, singers continually adjusted their internal models as the FAF manipulations became larger and smaller across trials. Alternatively, nonsingers appeared to search for the target note by initiating vocal productions at frequencies below the auditory target (0 cents) and increased their F_0 until they matched the note. Although nonsingers slightly adjusted their starting point, we did not find consistent evidence to suggest that sensorimotor adaptation occurred during FAF trials.

Multiple Internal Models for Vocal Control

Based on the findings presented in Chapter 1, it was believed that singers' internal models are more sensitive to subtle changes in auditory feedback, which was evident in their ability to compensate for subtle pitch shift manipulations.

This eventually led to the idea for Chapter 2 that multiple internal models may exist for vocal control. Recent evidence from the motor control of arm reaching supported this hypothesis (Donchin et al. 2003; Kalenscher et al. 2003; Osu et al. 2004; Wada et al. 2003; Wainscott et al. 2005; Wolpert and Kawato 1998). When applied to singing it was thought that trained singers should be able to modify internal representations that are specific to a particular target note. In order to assess this we presented singers with a sequence of three target notes, but they were only required to reproduce them one at a time when prompted by a contextual cue (the target note). A contextual cue was previously shown to be crucial in participants ability to acquire and switch between multiple internal models for arm reaching movements (Osu et al. 2004; Wada et al. 2003; Wainscott et al. 2005). This was mainly because of the tremendous difficulty often associated with learning to produce specific movements while using a manipulandum (robotic arm). For example, Karniel and Mussa-Ivaldi (2002) did not find evidence to suggest that participants could acquire and switch between internal models for two alternating viscous force fields, even after participants performed these movements in 4 sessions over 4 days. In the current study, the target notes served as a cue to inform what pitch to sing, similar to how colour was used to inform participants which manipulation they would experience on a given trial.

The results presented in Chapter 2 demonstrated how vocal control appears to be represented by multiple internal models. Results also suggested that singers' acoustic-motor mappings are capable of independent sensorimotor

recalibration that was observable within 50 ms of vocal onset. Of particular interest was the fact that aftereffects were target specific and did not generalize to other notes during the presentation of FAF. On the other hand, during test trials, following the conclusion of FAF, when feedback was suddenly returned to normal aftereffects were observed (in the 50 ms data) for the unaltered target. Transferred aftereffects (generalization) to a target that was not subjected to modified feedback has been observed in previous FAF (Jones & Keough 2008; Jones & Munhall 2005) and arm reaching investigations (Ghahramani et al. 1996; Shadmehr & Mussa-Ivaldi 1994). In Chapter 2 and other work (Jones & Keough 2008; Jones & Munhall 2005) pitch-shift manipulations were gradually presented during FAF. When feedback returned to normal, participants heard their F_0 1 semitone different than it was on the previous trial. Thus, the single-trial aftereffects observed in the median 50 ms F_0 data during the test trials for the unaltered pitch target may have been the result of the sudden and large changes in auditory feedback following training.

A unique aspect of studying voice F_0 while singing is that is not necessary (or actually possible) to require participants to initiate motor commands from a consistent starting point (a particular pitch). In the case of arm reaching studies, participants have been required to initiate movements from a fixed location (Imamizu & Kawato 2008; Krakauer et al. 1999; Osu et al. 2004; Wada et al. 2003). This has permitted the examination of feed-forward internal models within 250 ms of movement initiation (e.g., Wainscott et al., 2005), which has been argued to be a period of time where motor commands (trajectories) are little

influenced by closed-loop control or 'online' feedback. However, the data from Chapters 1 and 2 indicated that sensorimotor adaptation during a singing task could be observed within 50 ms of vocal onset. An advantage of this analysis was that it could account for the lack of aftereffects in participants' average productions (1500 ms data) during testing. We found that singers initiated vocal productions during test trials as if they were anticipating FAF and when they identified that their feedback was returned to normal they rapidly adjusted the pitch of their voice to levels indistinguishable from baseline.

Previous research has demonstrated that if the contextual information is ambiguous or not present at all and if the multiple environments are difficult to discriminate (Brashers-Krug and Shadmehr 1996; Gandolfo et al. 1996; Karniel and Mussa-Ivaldi 2002; Krakauer et al. 1999; Tong et al. 2002), then acquiring or switching between multiple internal representations is difficult (Wada et al. 2003). Following Experiment 1 in Chapter 2 we were uncertain whether singers were relying on the cue to modify their vocal productions or whether they were relying on the predictability of pitch shift sequences. Therefore, we removed the cue by having singers reproduce one target, which was present on each trial, while they received the same pitch shift manipulations as in Experiment 1. Eliminating the reliance on a contextual cue resulted in a similar pattern of compensatory responding as was observed in Experiment 1, however trial-by-trial adaptation was also found that was limited to the unaltered target note in the sequence.

In sum, the pattern of sensorimotor recalibration found in Chapter 2 seems to suggest that producing multiple target notes while singing requires participants

to employ unique motor commands to maintain accurate pitch control. The data indicated that trained singers consistently produced vocal pitch near the desired target frequencies, even in the presence of FAF. Overall, sensorimotor adaptation appears to be automatic, it is observable within 50 ms of vocal onset during training, and is uniquely linked with the motor commands for specific musical targets.

Ignoring Frequency-Altered Feedback: The Role of Task Instructions

In Jones and Keough (2008), it was argued that when trained singers detect large incongruities between perception and production they rely more on their internal model to produce the target. This implies that they may be capable of ignoring the error detected in auditory feedback and can rely on alternative mechanisms of their internal models to maintain accurate F_0 control. Indeed, this raises an interesting question about singers' internal models. For instance, were the early compensatory responses observed to subtle shifts in auditory feedback (Chapters 1 & 2) a result of singers attributing the perceived error to something internal (themselves), whereas the reduced compensation to large pitch manipulations (Jones & Keough, 2008) resulted from singers attributing the error to something external (experimenter)? Indeed, Kluzik and colleagues (2008) suggested that the brain may rely on the size of the error to determine how to assign 'blame' for the resulting discrepancy between the expected and actual sensory feedback. An alternative explanation may be that cognitive awareness of the altered feedback influenced singers' performance.

Related to these points are the results obtained by Zarate and Zatorre (2008). The authors had singers and nonsingers sing a musical scale while instructing them to either compensate or ignore the FAF. Zarate and Zatorre (2008) found that singers could almost completely ignore FAF, that is, not compensate when instructed to do so. On the other hand, nonsingers were unable to ignore the feedback as their F_0 responses indicated partial compensation. These findings suggest that singers' internal models may be under volitional control, whereas compensation is automatic in musically untrained participants. These data were consistent with Finney and Palmer's (2003) study examining the role of auditory feedback of musical performance in trained pianists. They found that completely removing pianist's auditory feedback while reciting well-rehearsed pieces from memory did not influence the quality of their performance, relative to when feedback was provided (Finney & Palmer, 2003). In regards to singing, the participants in Zarate and Zatorre (2008) received FAF between 1000-1500 ms following vocal onset. It is possible that similar mechanisms are used to monitor auditory feedback when perturbations occur at either mid utterance or at vocal onset and if this is the case then singers should be able to ignore FAF regardless of when it is presented. For example there is evidence to suggest that compensatory responses are similar regardless of whether vocal perturbations occur during mid utterance or are presented just prior to vocal onset (Larson et al., 2001). Larson and colleagues (2001) used pitch-shifted auditory feedback manipulations in two timing conditions to investigate whether the internal reference used to guide vocal is fixed or variable.

The pitch shift manipulation during the onset condition was unexpectedly turned on shortly following vocal onset and remained on throughout the remainder of the vocal production. The pitch shift manipulation during the OFF condition was present at vocal onset and was unexpectedly turned off and remained off during the vocalization. In both instances the manipulations occurred between 500-2500 ms following vocal onset. Thus, participants either heard their voice randomly shifted by 25, 100, or 200 cents during this time frame or suddenly heard their voice return to normal from the aforementioned shift values.

Larson and colleagues (2001) noted similar compensatory levels when participants initiated vocal productions with modified auditory feedback that was randomly removed during the vocal production or when they heard their feedback unexpectedly shifted during the utterance. The authors argued that in the absence of an absolute external reference that participants rely on an internal variable reference (Larson et al., 2001). Interestingly, when Hawco and Jones (2009) had participants produce an absolute target while auditory feedback was perturbed during mid utterance and they also observed similar compensatory responses, suggesting that producing a stable voice F_0 is based on the current, pre-manipulated F_0 value. Thus, it has been argued that in speech production the goal of the compensatory response is to overcome unintentional changes in voice F_0 during vocal production (Hawco & Jones, 2009). Alternatively, presenting participants with an absolute reference on each trial while manipulating auditory feedback prior to vocal onset was shown to elicit larger compensatory responses than shifting voice F_0 during mid utterance (Hawco &

Jones, 2009). This suggests that different mechanisms may be used to monitor auditory feedback at vocal onset than during mid utterance.

The idea that different systems may monitor auditory feedback may account for trained singers ability to largely suppress compensatory responding when pitch manipulations occurred between 1000-1500 ms following vocal onset in Zarate and Zatorre (2008). One possibility is that singers' internal models detected a violation between the efference copy and the actual sensory feedback. Recall that internal models are hypothesized to compare sensory feedback with motor acts by means of a comparator examining differences between perception and production. These differences are believed to be computed based on a corollary discharge, such that the output of an internal model maps the motor commands (e.g., efference copy) with the expected sensory feedback from the actions. When a match exists between perception and production the result is a net cancellation of the sensory input, which in turn causes a dampened sensory experience. Conversely, when there is a discrepancy between the perception and production of a motor act there is no corollary discharge to cancel the sensory feedback. As a consequence, there is an intensification of the sensory experience that potentially alerts us to environmental events (Heinks-Maldonado et al., 2005).

Unfortunately, explaining the results of Zarate and Zatorre (2008) in terms of an efference copy violation is not ideal, as the evidence suggests an efference copy violation is related to errors detected at vocal onset (Curio et al., 2000, Houde et al., 2002; Heinks-Maldonado et al., 2005, 2006). However, when mid

utterance perturbations were presented to participants in Hawco et al. (2009) a mismatch negativity (MMN) was found, rather than an N100. The differences in perturbation onset between Heinks-Maldonado et al. (2005) and Hawco and colleagues (2009) indicated that auditory feedback during both instances may be monitored by different underlying mechanisms. Indeed, a more likely explanation is that singers in Zarate and Zatorre (2008) attempted to maintain a stable F_0 by relying on alternative components (e.g., muscle memory, kinesthetic feedback) of their internal model to suppress compensatory responding, whereas nonsingers continually use auditory feedback to maintain their voice F_0 . This is consistent with the idea that internal models for the control of voice F_0 while singing are more entrenched in singers than nonsingers, which was previously addressed in Jones and Keough (2008) and in Chapters 1 and 2 of this dissertation.

Chapter 3 was designed to investigate whether singers and nonsingers are capable of voluntarily suppressing compensatory responses and sensorimotor adaptation to FAF when presented with gradual (-2 cent increments down to -100 cents) and constant (-100 cents) manipulations coinciding with vocal onset. Results indicated that singers and nonsingers were unable to suppress compensatory responding when presented with gradual or constant pitch manipulations. That is, the pattern of compensatory responding obtained when participants were asked to 'compensate' for FAF was indistinguishable from that observed when they were asked to 'ignore' FAF. Additionally, participants' median 50 ms F_0 values suggested that the level of sensorimotor

adaptation that occurred during the ignore condition was similar to the adaptation observed during the compensation condition.

The finding that compensatory responding is not easily suppressed by instructions to ignore feedback is consistent with previous studies using FAF (Hain et al., 2000), formant frequency manipulations (Munhall et al., 2009), and masking noise (Pick et al., 1989). Indeed, a recent study by Munhall and colleagues (2009), found that participants rapidly compensated for formant frequency manipulations when they were instructed to ignore the modified feedback. They also observed aftereffects when the manipulations were removed. Munhall and colleagues (2009) argued that compensatory responses to vowel modifications are not intentional strategic responses to the detection of auditory feedback manipulations. This is also congruent with the findings from the data in Chapter 3; whether participants were given explicit knowledge of the modified feedback does not seem to influence the pattern of responding during altered feedback trials. Moreover, it is possible that compensatory responses to modified feedback coinciding with vocal onset are related to the violation of an efference copy. Thus, once the efference copy has been violated, or an error has been detected, participants' internal models are automatically recalibrated and compensatory responses are initiated in an attempt to offset the auditory feedback manipulations, regardless of whether they were 'consciously' aware of the FAF. Conversely, presenting the pitch manipulations so they occur later into vocal production (1000-1500 ms in Zarate & Zatorre, 2008) may result in the

singer's ability to use the pre-shifted F_0 as a reference to guide the current production and suppress compensatory responding.

Recap and Future Directions on the Role of Auditory Feedback for Vocal Control

The primary goal of this dissertation was to address the importance of auditory feedback on the internal representations for F_0 control while singing. The data presented in this dissertation demonstrated that both singers and nonsingers use auditory feedback in the development and maintenance of internal models for vocal control. It was also shown that participants' internal models, like those for speech production, were capable of being recalibrated in the presence of very small (2 cent incremental shifts) or large (100 cent) changes in auditory feedback. An interesting result in Chapters 1 and 2 was that generalization was not observed in participants mean vocal productions following FAF trials. Indeed, a hallmark of internal models is that learning can be generalized to additional movements other than the one acquired through training. Recall that in Jones and Keough (2008) we found aftereffects that generalized to another note that singers did not produce with altered feedback. Moreover, generalization persisted across all blocks of test trials following FAF in trained singers. The lack of aftereffects following training with FAF led us to examine productions early after vocal onset (within 50 ms). This alternative assessment of sensorimotor adaptation allowed us to not only determine that the recalibration of internal models could be observed online during FAF trials (Chapter 1), but that adaptation to multiple internal models could be observed (Chapter 2).

Lastly, Chapter 3 identified that sensorimotor adaptation and compensation is automatic and occurs even when participants are consciously aware of the manipulation. Chapter 3 also demonstrated that musical training does not seem to influence responding to FAF, providing the perturbations coincide with vocal onset. Alternatively, if the modified feedback occurs during mid utterance then musical training appears to assist trained singers in the ability to suppress compensatory responses (Zarate & Zatorre, 2008). It may be the case that training for some singers' focuses on the kinesthetic feedback associated with musical sounds. Indeed, this would be particularly useful in the event that auditory feedback becomes reduced, as in singing in a choir. It should be noted that nonreflexive components (e.g., kinesthetic feedback) aside from auditory feedback have been shown to be influential in achieving precise vocal control. Murbe et al. (2004) and Larson et al. (2008) have also demonstrated that kinesthesia substantially contributes to singers pitch control, but very early in responses (< 100 ms). After 100 ms auditory feedback participates in F_0 control (Larson et al., 2008). On the other hand, an interesting result presented by Munhall and colleagues (2009) was that instructing participants to rely on the kinesthetic properties for F_0 control was insufficient to suppress compensatory responding to formant frequency manipulations. However, if the participants in Munhall et al. (2009) were not musically trained then it may be that they were unable to utilize the kinesthetic feedback to suppress compensatory responses. Moreover, given that we have argued nonsingers rely heavily on auditory feedback and that Munhall et al. (2009) did not eliminate it during their study, it is

likely that when auditory feedback is present nonsingers will use it to regulate the pitch of their voice over other sources of feedback.

Based on the research I have conducted thus far some questions come to mind that bare consideration for future research. For instance, at what point does the brain assign the error received during the gradual presentation of FAF to something external? It is thought that compensation to small pitch modifications could be attributable to oneself, however as the manipulations increase/decrease is there a point that participants' internal models assign the error to the experimenter? It is believed that a component of an internal model is dedicated for the detection of error and once it has been detected compensatory responses are elicited. Indeed, whether one is speaking or singing may also influence when compensatory responses are initiated. In English speech an individual's F_0 is a relative pitch target. Thus, it would be expected that the internal reference in speech may be less sensitive to perturbations than internal representations for singing. This may account for Natke and colleagues (2003) result that singing elicited greater compensatory responding than speech. I believe addressing the similarities and differences between speech and singing will identify how the brain monitors auditory feedback during each task respectively. It seems that during speech the goal is to maintain a stable voice F_0 around a relative frequency, whereas in singing the goal is to match a specific pitch target that may result in greater compensation.

I am also interested in investigating the time course of internal model acquisition in typically developing children. To put this in context, recall that a

hallmark of internal models is that generalization can transfer to additional movements other than the one acquired through training. Moreover, it seems that musical training can have an influence on vocal control while singing (Zarate & Zatorre, 2008). Interestingly, formal music training has been shown to have positive effects ('transfer') on other aspects of children's development, such as intelligence quotient (IQ), mathematic, verbal, and spatial abilities (Forgeard, Winner, Norton, & Schlaug, 2008; Rauscher et al., 1997; Schellenberg, 2004; Vaughn, 2000, but see Hyde et al., 2009). More recently, formal music training has also been found to influence cortical plasticity at a young age during development (Hyde et al.; Moreno et al., 2009), with the benefits of training still evident in adulthood (Wong et al., 2007). Hyde et al. (2009) and Moreno et al. (2009) were the first to examine structural brain and behavioural changes during development following long-term music training. Both studies found no differences between children prior to the commencement of the experiment. Thus, it was argued that any performance increase found in musically trained children compared to controls was due to training. Although the contribution of a genetic predisposition cannot be ruled out completely, the findings do support the view that the brain differences and enhanced behavioural abilities (e.g., reading and pitch discrimination in speech) found in these studies and adults is due to intensive music training (Hyde et al.; Moreno et al.; Wong et al., 2007).

Indeed, a question arising from these findings is the contribution of genetics (predisposition) and nurture on music ability and neural plasticity. Moreno et al. (2009) found that musical training led to an increase in reading and

pitch discrimination abilities in speech. However, what is surprising about the results of Moreno et al. (2009) was that participants only had 6 months of musical training. Moreover, genetic differences were ruled out because the authors did not find statistical differences between the musical and painting groups prior to training. Also, the parents of the children filled out questionnaires to confirm that none had formal music or painting training. It was argued that the enhanced ERP components corresponding to reading and pitch discrimination abilities in speech occurred as a result of music training. Thus, the developmental benefits from music training extend to reading, highlighting that a short period of training is sufficient to demonstrate neural plasticity on the functional organization of children's brains.

On the other hand, it is important to note that no improvements were observed for performance IQ, for WISC indexes (POI & VCI), and for verbal memory in the Moreno et al. study (2009). It is possible though that general maturational and development effects in children could account for the lack of differences. Although transfer advantages with music training have been shown, not all studies have found congruous results. For example, Hyde et al. (2009) used functional magnetic resonance imaging (fMRI) to measure structural brain changes in the developing brain following long-term (15 months) music training. Following 15 months of music training children did not show increased performance in visual-spatial and verbal transfer domain outcomes over those children who did not receive formal lessons (Hyde et al., 2009).

Despite this result, Hyde and colleagues (2009) were the first group to examine brain and behavioural changes during development prior to and after long-term music training. Interestingly, at baseline there were no structural brain differences found between participants in the instrumental group and those in the control group. At the conclusion of training, structural changes were observed in motor and auditory areas associated with music training that was correlated with improvements on motor and auditory-musical tests (Hyde et al., 2009). The results of Hyde et al. (2009) and Moreno et al. (2009) cannot completely rule out whether those children in the music training condition had a genetic predisposition for musicality or not. However, their data suggests that structural brain differences following training were due to music training. The reason for the lack of transfer effects reported in Hyde et al. (2009) may be a result of the experience the children received during the control condition. Control participants were subjected to weekly 40 min group music classes at school, which consisted of singing and playing percussion and bells (Hyde et al., 2009). Arguably, this brief and consistent exposure to music may account for the lack of differences exhibited on 'other' non-musical tasks following the 15-month period. Alternatively, the children in Hyde et al. (2009) were between the ages of 5-7 years whereas those in the Moreno et al. were 8 years of age. Thus, one cannot rule out age as a potential factor influencing transfer effects during development. This could be assessed in future studies examining long-term music training. Overall, the evidence seems to suggest that music training does indeed have

positive effects that generalize to positively influence other cognitive aspects during development.

The importance of investigating internal representations during development in children may help to identify whether positive transfer effects occur within a window corresponding to internal model acquisition. For instance, no transfer was observed in children 5-7 years of age (Hyde et al., 2009), whereas children 8 years old did profit from music training (Moreno et al., 2009). Overall, determining when internal models are acquired and when transfer effects occur could be particularly relevant for developmental researchers; in that children having difficulty with reading and mathematics, for example, might benefit from music training. Indeed, music is ubiquitous across every known culture, regardless of technological sophistication. Given the strong enculturation of music (Hannon & Trainor, 2007), it seems reasonable to assume that children would enjoy learning to make music and if the positive transfer effects assist the majority of those in training then it could be of valuable assistance to programs designed to improve developmental deficits.

Conclusion

We have shown that whether we manipulate auditory feedback in large increments or gradually, whether it is for a single music note or multiple notes, or whether it is the instructions associated with the task (ignore or compensate), the result is a change in the sensorimotor representations (internal models) underlying voice F_0 control while singing. One goal has been to demonstrate the

plasticity of the nervous system by examining how changes in singers and nonsingers auditory feedback can alter target specific internal representations. Indeed, producing multiple target notes while singing was shown to require participants to employ unique motor commands for each target. We have also argued that although the human voice has the potential to be initiated at unpredictable frequencies during onset, trained singers consistently produced vocal pitch at frequencies near the desired target, even in the presence of FAF. Overall, it appears that compensation and adaptation to FAF are automatic and are influenced little by 'conscious' control, and are uniquely associated with the motor commands for specific musical targets.

References

- Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude. *J. Acoust. Soc. Am.* 119, 2363-2371.
- Blakemore, S. J., Wolpert, D., and Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, 11, R11-16.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5, 341-345.
- Bosshardt, H. G., Sappok, C., Knipschild, M., and Holscher, C. (1997). Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of Psycholinguistic Research*, 26, 425-448.
- Borden, G. J. (1979). An interpretation of research on feedback interruption in speech. *Brain and Language* 7, 307-319.
- Brashers-Krug, T., Shadmehr, R., and Bizzi, E. (1996). Consolidation in human motor memory. *Nature* 382, 252-255.
- Burnett, T.A., Freedland, M.B., Larson, C.R., and Hain, T.C. (1998). Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* 103, 3153-3156.
- Burnett, T.A. and 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.
- Burnett, T. A., McCurdy, K. E., and Bright, J. C. (2008). Reflexive and volitional voice fundamental frequency responses to an anticipated feedback pitch error. *Exp. Brain Res.* 191, 341-351.
- Burnett, T. A., Senner, J. E., and Larson, C. R. (1997). Voice F0 responses to pitch-shifted auditory feedback: a preliminary study. *J. Voice* 11, 202-211.
- Cowie, R., and Douglas-Cowie, E. (1992). Postlingually acquired deafness. *Trends in linguistics, studies and monographs*. Mouton de Gruyter, New York.
- Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cog. Sci.* 4, 423-431.

- Donath T. M., Natke U., and Kalveram K. T. (2002). Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *J. Acoust. Soc. Am.* 111, 357-366.
- Donchin, O., Francis, J. T., and Shadmehr, R. (2003). Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: Theory and experiments in human motor control. *J. Neurosci.* 23, 9032-9045.
- Eliades, S. J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, 453, 1102-1106.
- Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *Acoust. Soc. Am.* 70, 45-50.
- Fairbanks, G. (1954). Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem. *J. Speech Hear. Dis.* 19, 133-139.
- Finney, S., and Palmer, C. (2003). Auditory feedback and memory for music performance: Sound evidence for an encoding effect. *Mem. Cog.* 31, 51-64.
- Flanagan, J. R., and Wing, A. M. (1993). Modulation of grip force with load force during point-to-point arm movements. *Exp. Brain Res.* 95, 131-143.
- Gandolfo, F., Mussa-Ivaldi, F. A., and Bizzi, E. (1996). Motor learning by field approximation. *Proc. Natl. Acad. Sci.* 93, 3843-3846.
- Ghahramani, Z., and Wolpert, D. M. (1997). Modular decomposition in visuomotor learning. *Nature* 386, 392-395.
- Ghahramani, Z., Wolpert, D. M., and Jordan, M. I. (1996). Generalization to local remappings of the visuomotor coordinate transformation. *J. Neurosci.* 16, 7085-7096, 1996.
- Gidley Larson, J. C., Bastian, A. J., Donchin, O., Shadmehr, R., and Mostofsky, S. H. (2008). Acquisition of internal models of motor tasks in children with autism. *Brain*, 131, 2894-2903.
- Gomi, H., Shidara, M., Takemura, A., Inoue, Y., Kawano, K., and Kawato, M. (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys I. Simple Spikes. *Journal of Neurophysiology*, 80, 818-831.

- Grell, A., Sundberg, J., Ternstrom, S., Ptok, M., and Altenmuller, E. (2009). Rapid pitch correction in choir singers. *J. Acoust. Soc. Am.* 126, 407-413.
- Guenther, F. H., and Perkell, J. S. (2004). A Neural Model of Speech Production and Its Application to Studies of the Role of Auditory Feedback in Speech. In B. Maassen, R. Kent, H. Peters, P. Van Lieshout and W. Hulstijn (Eds.), *Speech Motor Control in Normal and Disordered Speech* (Eds.) (pp. 29-49). Oxford: Oxford University Press.
- Hafke, H. Z. (2008). Nonconscious control of fundamental voice frequency. *J. Acoust. Soc. Am.* 123, 273-278.
- Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and 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.
- Hawco, C. S., and Jones, J. A. (2009). Control of vocalization at utterance onset and mid-utterance: different mechanisms for different goals. *Brain Res.* 1276, 131-139.
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., and Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophys.* 42, 180-190.
- Heinks-Maldonado, T. H., Nagarajan, S. S., and Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport* 17, 1375-1379.
- Held, R. (1965) Plasticity in sensory-motor systems. *Scientific American*, 213, 84-94.
- Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Sci.* 279, 1213-1216.
- Imamizu, H., and Kawato, M. (2008). Neural correlates of predictive and postdictive switching mechanisms for internal models. *J. Neurosci.* 28, 10751-10765.
- Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R. Putz, B., et al. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, 403, 192-195.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9, 304-313.
- Jones, J. A., and Keough, D. (2008). Auditory-motor mapping for pitch control

- insingers and nonsingers. *Exp. Brain Res.* 190, 279-287.
- Jones, J. A., and Munhall, K. G. (2000). Perceptual calibration of F0 production: evidence from feedback perturbation. *J. Acoust. Soc. Am.* 108, 1246-1251.
- Jones, J. A., and Munhall, K. G. (2002). The role of auditory feedback during phonation: Studies of Mandarin tone production. *J. Phonetics.* 30, 303-320.
- Jones, J. A., and Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Current Bio.* 15, 1768-1772.
- Kalenscher, T., Kalveram, K. T., and Konczak, J. (2003). Effects of two different dynamic environments on force adaptation: Exposure to a new force but not the preceding force experience accounts for transition and after-effects. *Motor Control* 7, 242-263.
- Karniel, A., and Mussa-Ivaldi, A. (2002). Does the motor control system use multiple models and context switching to cope with a variable environment? *Exp. Brain Res.* 143, 520-524.
- Kawahara, H. (1998). Hearing voice: Transformed auditory feedback effects on voice pitch control. In D. F. Rosenthal and H. G. Okuno (Eds.), *Computational auditory scene analysis* (pp. 335-349). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Kawahara, H., Masuda-Katsuse, I., and de Cheveigne, A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Comm.* 27, 187-207.
- Kawato, M., Furukawa, K., and Suzuki, R. (1987). A Hierarchical Neural-Network Model for Control and Learning of Voluntary Movement. *Biological Cybernetics*, 57, 169-185.
- Kawato, M., and Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, 68, 95-103.
- Keough, D., and Jones, J. A. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *J. Acoust. Soc. Am.* 126, 837-846.
- Kirchner, J. A., and Wyke, B. D. (1965). Articular reflex mechanisms in the larynx. *Annals Oto. Rhino. Laryngo.* 74, 749-768.

- Kitazawa, S., Kimura, T., and Yin, P. B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. *Nature*, 392, 494-497.
- Krakauer, J. W., Ghilardi, M. F., and Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neurosci.* 2, 1026-1031.
- Krakauer, J. W., and Shadmehr, R. (2007). Towards a computational neuropsychology of action. In: *Computational Neuroscience: Theoretical Insights into Brain Function*, 165. Elsevier.
- Lane, H., and Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *J. Speech Hear. Res.* 14, 677-709.
- Lane, H., and Webster, J. (1991). Speech deterioration in postlingually deafened adults. *J. Acoust. Soc. Am.* 89, 859-866.
- Larson, C. 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, C. R., Altman, K. W., Liu, H., and Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice F0 control. *Exp. Brain Res.* 187, 613-621.
- Larson, C. R., Burnett, T. A., Bauer, J. J., and Kiran, S. (2001). Comparison of voice F0 responses to pitch-shift onset and offset conditions. *J. Acoust. Soc. Am.* 110, 2845-2848.
- Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (2000). Effects of pitch-shift velocity on voice F0 responses. *J. Acoust. Soc. Am.* 107, 559-564.
- Lee, B. S. (1950). Effects of Delayed Speech Feedback. *J. Acoust. Soc. Am.* 22, 824-826.
- Liu, H., and Larson, C. R. (2007). Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *Journal of the Acoustical Society of America*, 122, 3671-3677.
- Liu, H., Zhang, Q., Xu, Y., and Larson, C.R. (2007). Compensatory responses to loudness-shifted voice feedback during production of Mandarin speech. *Journal of the Acoustical Society of America*, 122, 2405-2412.
- Loui, P., Guenther, F. H., Mathys, C., and Schlaug, G. (2008). Action-perception

- mismatch in tone-deafness. *Current Bio.* 18, 331-332.
- Munhall, K. G., MacDonald, E. N., Byrne, S. K., and Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *J. Acoust. Soc. Am.* 125, 384-390.
- Murbe, D., Pabst, F., Hofmann, G., Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback: A longitudinal study of singers' pitch control. *J. Voice* 18, 236-241.
- Natke, U., Donath, T. M., and Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *J. Acoust. Soc. Am.* 113, 1587-1593.
- Nowak, D. A., Topka, H., Timmann, D., Boecker, H., and Hermsdörfer, J. (2007). The role of the cerebellum for predictive control of grasping. *Cerebellum*, 6, 7-17.
- Oller, D. K., and Eilers, R. E. (1988). The role of audition in infant babbling. *Child Dev.* 59, 441-449.
- Osu, R., Hirai, S., Yoshioka, T., and Kawato, M. (2004). Random presentation enables subjects to adapt to two opposing forces on the hand. *Nature* 7, 111-112.
- Pape, D., and Mooshammer, C. (2006). Is intrinsic pitch language dependent? Evidence from a cross-linguistic vowel pitch experiment (with additional screening of the listeners' DL for music and speech). In MULTILING-2006, paper 018.
- Peretz, I., and Hyde, K. (2003). What is specific to music processing? Insights from congenital amusia (tone deafness). *Trends Cog. Sci.* 7, 362-367.
- Pick, H. L., Jr., Siegel, G. M., Fox, P. W., Garber, S. R., and Kearney, J. K. (1989). Inhibiting the Lombard effect. *J. Acoust. Soc. Am.* 85, 894-900.
- Prochazka, A., Clarac, F., Loeb, G. E., Rothwell, J. C., and Wolpaw, J. R. (2000). What do reflex and voluntary mean? Modern views on an ancient debate. *Experimental Brain Research*, 130, 417-432.
- Proteau, L., Marteniuk, R. G., Girouard, Y., and Dugas, C. (1987). On the type of information used to control and learn an aiming movement after moderate and extensive training. *Human Mov. Sci.* 6, 181-199.
- Proteau, L., Marteniuk, R. G., and Levesque, L. (1992). A sensorimotor basis for

- motor learning: Evidence indicating specificity of practice. *Quart. J. Exp. Psych.* 44A, 557-575.
- Purcell, D. W., and Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *J. Acoust. Soc. Am.* 119, 2288-2297.
- Purcell, D. W., and Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *J. Acoust. Soc. Am.* 120, 966-977.
- Sainburg, R. L., Ghez, C., and Kalakanis, D. (1999). Intersegmental dynamics are controlled by sequential anticipatory, error correction, and postural mechanisms. *J. Neurophys.* 81, 1045-1056.
- Sapir, S., McClean, M. D., and Larson, C. R. (1983). Human laryngeal responses to auditory stimulation. *J. Acoust. Soc. Am.* 73, 315-321.
- Shadmehr, R., and Moussavi, Z. M. K. (2000). Spatial generalization from learning dynamics of reaching movements. *J. Neurosci.* 20, 7807-7815.
- Shadmehr, R., and Mussa-Ivaldi, F. A. (1994). Adaptive representation of Dynamics during learning of a motor task. *J. Neurosci.* 14, 3208-3224.
- Shaiman, S., and Gracco, V. L. (2002). Task-specific sensorimotor interactions in speech production. *Exp. Brain Res.* 146, 411-418.
- Schultz-Coulton, H. J. (1978). The neuromuscular phonatory control system and vocal function. *Acta Otolaryngology* 86, 142-153.
- Smotherman, M., Zhang, S., and Metzner, W. (2003). A neural basis for auditory feedback control of vocal pitch. *J. Neurosci.* 23, 1464-1477.
- Sundberg, J. (1987). *The science of the singing voice*. Northern Illinois University Press, Dekalb, IL.
- Tong, C., Wolpert, D. M., and Flanagan, J. R. (2002). Kinematics and dynamics are not represented independently in motor working memory: Evidence from an interference study. *J. Neurosci.* 22, 1108-1113.
- Toyomura, A., Koyama, S., Miyamaoto, T., Terao, A., Omori, T., Murohashi, H., and Kuriki, S. (2007). Neural correlates of auditory feedback control in human. *Neurosci.* 146, 499-503.
- Tramblay, S., Houle, G., and Ostry, D. J. (2008). Specificity of speech motor learning. *J. Neurosci.* 28, 2426-2434.

- Wada, Y., Kawabata, Y., Kotosaka, S., Yamamoto, K., Kitazawa, S., and Kawato, M. (2003). Acquisition and contextual switching of multiple internal models for different viscous force fields. *Neurosci. Res.* 46, 319-331.
- Wainscott, S. K., Donchin, O., and Shadmehr, R. (2005). Internal models and contextual cues: Encoding serial order and direction of movement. *J. Neurophysio.* 93, 786-800.
- Ward, D., and Burns, E. (1978). Singing without auditory feedback. *J. Res. in Singing* 1, 4-44.
- Watts, C., Murphy, J., and Barnes-Burroughs, K. (2003). Pitch matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with nontalented singing voices in conditions of varying feedback. *Journal of Voice*, 17, 185–194.
- Wolpert, D. M., and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks* 11, 1317-1329.
- Wyke, B. D. (1974). Laryngeal neuromuscular control systems in singing. A review of current concepts. *Folia Phoniatria* 26, 295-306.
- Yates, A. J. (1963). Delayed auditory feedback. *Psych. Bulletin* 60, 213–251.
- Yoshida, Y., Saito, T., Tanaka, Y., Hirano, M., Morimoto, M., and Kanaseki, T. (1989). Laryngeal sensory innervation: Origins of sensory nerve fibers in the nodose ganglion of the cat. *J. Voice* 3, 314-320.
- Zarate, J. M., and Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage* 40, 1871-1887.
- Zarate, J. M., and Zatorre, R. J. (2005). Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Annals New York Acad. of Sci.* 1060, 404-408.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends Cog. Sci.* 6, 37-46.