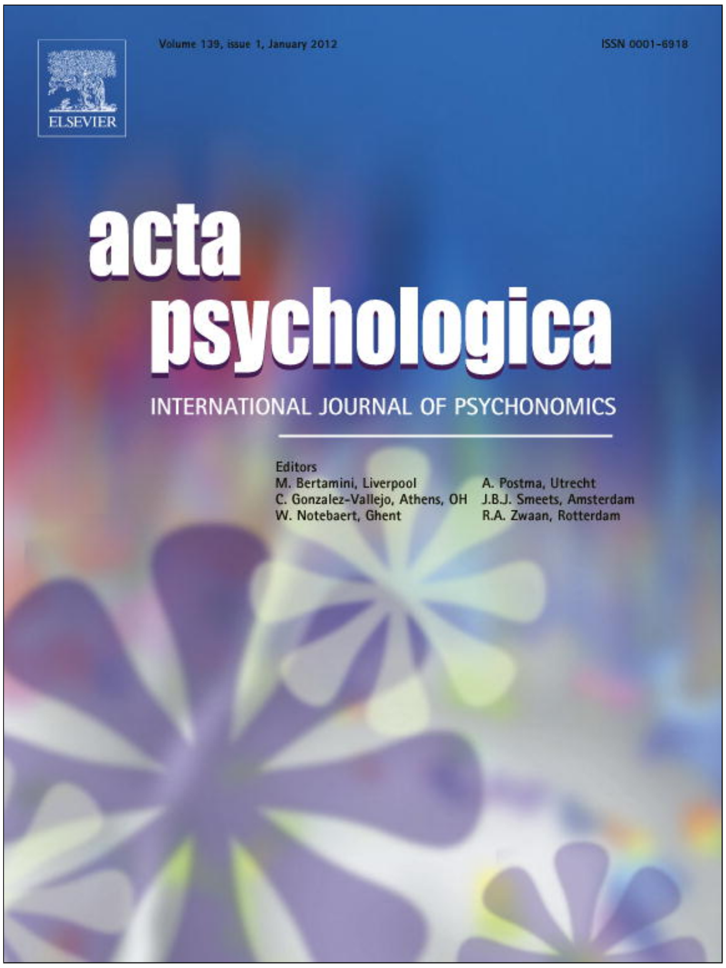
Volume 139, issue 1, January 2012
ISSN 0001-6918
ELSEVIER
acta psychologica
INTERNATIONAL JOURNAL OF PSYCHONOMICS
Editors
M. Bertamini, Liverpool
C. Gonzalez-Vallejo, Athens, OH
W. Notebaert, Ghent
A. Postma, Utrecht
J.B.J. Smeets, Amsterdam
R.A. Zwaan, Rotterdam

Acta Psychologica 139 (2012) 166–177

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

# Effects of altered auditory feedback across effector systems: Production of melodies by keyboard and singing

Peter Q. Pfordresher *, James T. Mantell

*University at Buffalo, State University of New York, United States*

ARTICLE INFO

*Article history:*
Received 28 February 2011
Received in revised form 25 July 2011
Accepted 20 October 2011
Available online 17 November 2011

*PsychINFO codes:*
2340 (motor processses)
2326 (auditory and speech)

*Keywords:*
Altered auditory feedback
Delayed auditory feedback
Music performance
Singing

ABSTRACT

We report an experiment that tested whether effects of altered auditory feedback (AAF) during piano performance differ from its effects during singing. These effector systems differ with respect to the mapping between motor gestures and pitch content of auditory feedback. Whereas this action-effect mapping is highly reliable during phonation in any vocal motor task (singing or speaking), mapping between finger movements and pitch occurs only in limited situations, such as piano playing. Effects of AAF in both tasks replicated results previously found for keyboard performance (Pfordresher, 2003), in that asynchronous (delayed) feedback slowed timing whereas alterations to feedback pitch increased error rates, and the effect of asynchronous feedback was similar in magnitude across tasks. However, manipulations of feedback pitch had larger effects on singing than on keyboard production, suggesting effector-specific differences in sensitivity to action-effect mapping with respect to feedback content. These results support the view that disruption from AAF is based on abstract, effector independent, response–effect associations but that the strength of associations differs across effector systems.



## 1. Introduction

When individuals produce a sequence of motor actions across a span of time, these actions are accompanied by perceptual events that result from actions. In certain domains, such as speaking and music performance, these perceptual consequences constitute goals for actions, and auditory feedback thus provides information about whether the appropriate goal has been met. Based in part on these observations, some have suggested that a close coupling exists between the mental representations used to plan actions and the representations used to monitor the consequences of these actions (e.g., Hommel, Muessler, Aschersleben & Prinz, 2001; MacKay, 1987; Prinz, Aschersleben, & Koch, 2009). If, as these theories suggest, perception and action share a common representation, then fluent production should depend on the coordination of perceptual feedback events with actions. This reliance is demonstrated by the disruptive effects of altered auditory feedback (AAF) during the production of speech and music (for reviews see Finney, 1999; Howell, 2004; Pfordresher, 2006; Yates, 1963). Disruptive effects of AAF differ from the effects of masking or removing auditory feedback, which have been found to yield negligible effects in musical keyboard production (Finney & Palmer, 2003; Pfordresher, 2005; Repp, 1999), and have been found to yield considerably smaller effects on singing than effects of AAF (Mürbe, Friedemann, Hofmann, & Sundberg, 2002, 2004; Ward & Burns, 1978).

The fact that production relies on sensorimotor coordination is interesting in itself, but perhaps a more compelling question is, how can AAF interference effects inform models of cognitive organization for perception and action? A related issue that has received differing support has to do with the role of different effector systems (i.e., the motor systems responsible for action production): whereas some research suggests that perception–action associations exist at an abstract level of representation that may extend across effector systems (e.g., Cohen, Ivry, & Keele, 1990; Grafton, Hazeltine, & Ivry, 1998; Howell, 2001; MacKay & Bowman, 1969; Palmer & Meyer, 2000), evidence also exists for reduced effects of sensorimotor interactions after switching response mode (Yamaguchi & Proctor, 2009). In addition, some neuroimaging evidence is consistent with the idea that perception–action associations are effector specific (e.g., Buccino et al., 2001). The research summarized in this paper addresses the degree to which effects of AAF generalize across two effector systems that are used to produce music: The hand–digit system (used for keyboard performance) and the vocal system (used for singing). Broadly speaking, effector independent effects of AAF suggest that disruption occurs at an abstract level of representation and are based on a general sensitivity to correlations between perception and action. By contrast, effector specific effects may reflect task-specific learned associations.

* Corresponding author at: Department of Psychology, 355 Park Hall, University at Buffalo, Buffalo, NY 14260, United States. Tel.: +1 716 645 0234; fax:+1 716 645 3801.
*E-mail address:* pqp@buffalo.edu (P.Q. Pfordresher).

0001-6918/$ – see front matter 
doi:10.1016/j.actpsy.2011.10.009

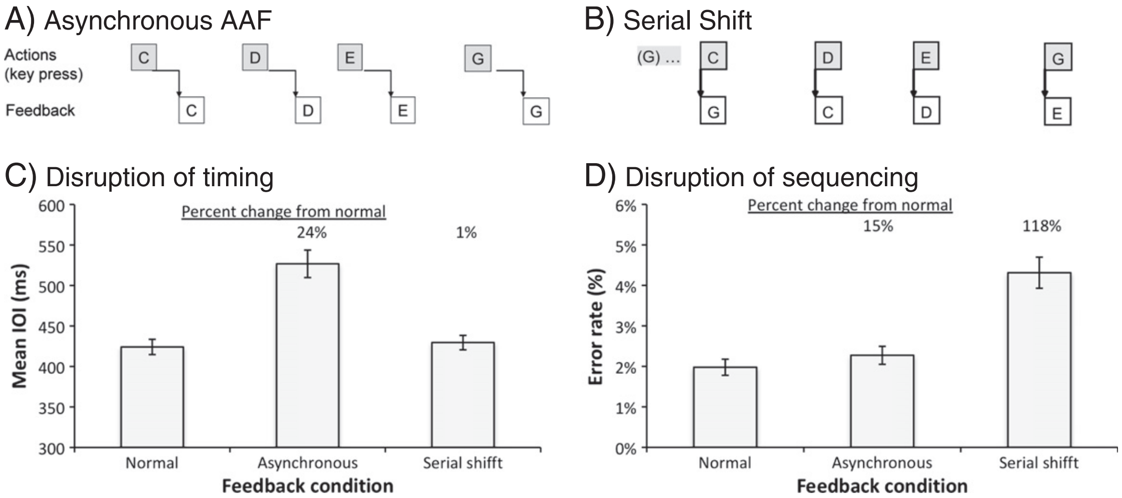


**Fig. 1.** Schematic illustrations of altered auditory feedback (AAF) manipulations that result in asynchronies (A) or alterations of contents (B) between perception and action. Gray boxes indicate the timing (left-to-right) and contents (letters) of produced actions. In this context, "contents" refers to a motor gesture (e.g., a piano key press) that under normal circumstances would lead to the pitch indicated by the letter. White boxes refer to timing and content of resulting perceptual events. Lower plots show pooled means (see text) across normal and AAF conditions that represent effects on timing (C) and the accuracy of sequencing (D). Error bars represent the between-participants standard error of the mean.

### *1.1. Dissociation of sequencing and timing effects from AAF*

The most well known form of AAF is delayed auditory feedback, in which a constant time lag is added to the onsets of perceptual feedback during production. Delays within the range of 100–400 ms can disrupt the performance of music on a keyboard, leading to slowed production (e.g., Finney, 1997; Gates, Bradshaw, & Nettleton, 1974), increased errors (e.g., Finney, 1997), and increases in timing variability (e.g., Pfordresher & Palmer, 2002). Similar effects have been found in speech (e.g., Black, 1951; Fairbanks & Guttman, 1958; Lee, 1950) and in other musical instruments (e.g., Havlicek, 1968).

More recent research on the effects of AAF during music production has sought to control the temporal coordination between actions and auditory feedback. Specifically, one can partition the effects of traditional delayed auditory feedback into two possible components: *feedback synchrony* and *feedback contents*. Illustrative schematic examples are shown in Fig. 1 (for further discussion, see Pfordresher, 2006; Pfordresher & Kulpa, 2011). A manipulation of feedback synchrony (Fig. 1A) causes feedback events (here, musical tones), to lag behind the actions associated with their production (e.g., a key press on a keyboard) yet occur before the next produced action. Importantly, in such circumstances an action is always followed by the anticipated event category (e.g., pressing the key for middle C leads to the associated pitch for middle C). Thus, the disruptive effects of such alterations can only be attributed to onset timing.[1]

A qualitatively different kind of AAF manipulation involves changing the contents of auditory feedback while maintaining synchronization between perception and action (or asynchronies must be too small to be noticed). In such cases, disruption must occur because the event category represented by auditory feedback differs from the anticipated event category. When participants experience serially shifted AAF, the onset time of a motor act (a key press on the keyboard, or the initial phonation of a syllable) coincides with a feedback event whose pitch matches a pitch from a different serial position in the sequence. The serial separation between actions and feedback events is kept constant; Fig. 1B shows a serial shift with a lag of 1, where each action produces the pitch associated with the previous serial position. Interestingly, other manipulations to feedback contents, such as presenting a randomly selected pitch or transposing the feedback melody, do not disrupt production (Finney, 1997; Pfordresher, 2005, 2008).

In one sense, AAF that is serially shifted by a lag of 1 is similar to asynchronous feedback, in that both manipulations present feedback information "too late." However, they differ in three critical respects. First, as described above, the fact that serial shifts are synchronous with actions means that disruption must be based on the mismatch between feedback contents and expected contents, whereas the same basis cannot be true of asynchronous feedback (as manipulated here). Second, the effects of these manipulations differ qualitatively (as described below). Finally, serial shifts that present future events (reported in Pfordresher & Palmer, 2006) lead to levels of disruption similar to that of serial shifts that present past events. Thus these manipulations of AAF differ in theoretically important qualitative respects.

Manipulations of feedback synchrony and feedback contents have qualitatively distinct effects on musical keyboard production, as shown in the lower part of Fig. 1. The illustrated data were pooled across three studies in which participants experienced normal feedback, asynchronous feedback (with delays equal to 33% of IOIs in a trial), and serial shifts of lag 1 (Benitez, 2005; Pfordresher, 2003, Experiment 4; Pfordresher et al., 2010). Participants in these experiments included pianists (with at least 8 years of formal training on the piano, N = 28) and non-pianists (N = 101) who performed short melodies from memory that were unfamiliar before learning. Both groups demonstrated the same pattern of results. Fig. 1C shows how asynchronous and serially shifted AAF influence production rate, measured by the mean of produced inter-onset intervals (IOIs, the time between successive key presses) in a trial. Importantly, participants were instructed to maintain a target tempo of 500 ms during AAF (though participants often exceed this rate when performing with normal feedback); higher mean IOIs represent greater slowing of performance timing. As can be seen, asynchronous feedback considerably slows production compared to normal feedback, whereas serial shifts have negligible effects on produced timing.[2]

[1] Traditional delayed auditory feedback using fixed delays often leads to asynchronous AAF. However, if the time of the delay is equal to the time between produced events, then delayed auditory feedback can lead to relationship between perception and action more like the serial shift of feedback (shown in Fig. 1B).

[2] The percent of change between AAF performance and normal feedback performance is calculated as $[((AAF - normal)/normal) \times 100]$. This simple measure of performance change allows comparisons of these pooled data with the data from the current study by controlling for differences in base (normal) levels of performance across participants (see Section 3).

Although effects of AAF on IOI are seen for asynchronous but not serially shifted feedback, the reverse is observed for measurements of error rates (incorrect key presses), which reflect the accuracy of action sequencing, shown in Fig. 1D. According to this measure, serial shifts yield a much more disruptive effect on note production than does asynchronous AAF. Taken together, measures of timing and accuracy demonstrate the *sequencing/timing dissociation* in the effect of AAF on music performance: asynchronous AAF disrupts timing (mean IOI) but not sequencing (note accuracy), whereas serially shifted AAF disrupts sequencing but not timing (Pfordresher, 2003, 2006; Pfordresher & Kulpa, 2011). The current work examines the generality of this dissociation across effector systems (see Section 1.2).

The patterns of performance across AAF manipulations have important implications for the way the representation of perception and action is conceptualized. With respect to the role of auditory feedback, the fact that alterations of contents (i.e., serial shifts) on their own can disrupt production (see Fig. 1D) suggests that the interfering effects of AAF can be based specifically on alterations to feedback content, thus resulting from conflicts between the planned outcomes of actions and actual auditory feedback. Moreover, the sequencing/timing dissociation may have important implications for the cognitive representation of perception and action. These effects suggest that the coordination of perception and action is guided by a common representation that is stratified with respect to time-scale. This claim follows from other research in motor control. Some researchers, using manual tapping paradigms, have suggested that event timing in production may be planned separately from serial order (Krampe, Mayr, & Kliegl, 2005; MacKay, 1987; Rosenbaum, Kenny, & Derr, 1983). With respect to the role of auditory feedback, this framework suggests that the mental representation of onset timing is distinct from the mental representation of serial order, both levels being used jointly to plan motor actions and to interpret perceptual feedback from those actions. According to this view (see Pfordresher, 2006 for details), a common representation of sequence structure (including sequencing and timing) is used both to plan produced events and to perceive the timing and content of feedback. During normal feedback, planned events are activated based on sequential content and timing, with feedback supplementing activation of the same events at the time of production. When AAF occurs, however, perceptual inputs add activation to time points (for asynchronous AAF) or event contents (for serial shifts) that conflict with the planned timing and/or contents of the current planned event. In other words, asynchronous and serially shifted AAF may selectively perturb two dissociable time-scales in a common representation used both to plan actions and to perceive auditory patterns (this hypothesis is more extensively examined in Pfordresher, 2006).

However, this dissociation has never been demonstrated in domains outside of melodic performance on a piano keyboard. As such, it is an open question as to whether these effects are constrained by effector system (manual) and/or domain (music). In fact, research concerning the effects of delayed auditory feedback on speech has suggested that feedback contents do not influence production, in that converting feedback to a non-speech signal does not alter the disruptive effect of delayed feedback to speech (e.g., Howell, 2007; Howell & Archer, 1984; Howell, Powell, & Khan, 1983; Howell & Sackin, 2002). One possible interpretation of these conflicting results is that serially shifted feedback disrupts action planning, whereas eliminating the meaningfulness of feedback (as in the studies of Howell et al.) does not (Pfordresher, 2006). However, the fact that these findings come from studies using different effector systems (hand/digit, versus vocal) representing different domains (music versus speech) makes it impossible to draw a firm conclusion (cf. Howell, 2004). Thus, we next consider the possibility that the role of auditory feedback differs across effector systems.

### *1.2. Perception–action associations in different effector systems*

Models of perception and action that inspired the account described above were predicated on the idea that the representations underlying perception, action, and planning function at an abstract level; in other words, the mental codes that represent perception and action are amodal (not system-specific). According to this view, action plans take the form of codes representing goals for actions, rather than muscle-specific commands, and perception–action associations are based on the degree of coherence between abstract codes used for action plans and similarly coded perceptual events (Hommel et al., 2001). Abstractionist views such as this clearly lead to the prediction that effects found for keyboard production should also be found for vocal production. However, recent neuroimaging evidence suggests that perception–action associations may differ across effector systems (Buccino et al., 2001), and behavioral evidence suggests that associations between perception and action in music may be heavily influenced by training within an effector system (Drost, Rieger, & Prinz, 2007). Beyond the empirical support, there are theoretical reasons to predict that effects of AAF on vocal production may differ from those found for keyboard production.

Howell et al. (1983) offered important reflections on possible differences in the vocal system and the manual production of keyboard sequences, in the context of evaluating the viability of "error monitoring" accounts for the role of auditory feedback (cf. Borden, 1979). In their paper, they referred specifically to differences across speech and non-speech tasks, although their comments clearly reflect a concern with the vocal system (producing either speech or a melody) versus production with other effector systems (such as the hand–digit system used for keyboard production):

> Consider music played on keyboard instruments. With electronic organs the pitch can be transposed at a flick of a switch without affecting performance. For a musician, it does not matter what the relation between an action and its outcome is; therefore, the feedback is not used to regulate the action. On the other hand, if a musician is required to regulate his or her meter with respect to other players, similar effects to those produced by DAF occur. (p. 773)

Two important predictions are suggested by this quote. First, the fact that perception–action relationships are at least somewhat arbitrary in (manual) music performance, according to this account, leads to a prediction regarding *susceptibility to disruption*. According to this logic, the motor gestures during vocalization (laryngeal tension) are more directly related to perceptual outputs (in particular, pitch) than are motor gestures in (non-vocal) musical performance tasks. As such, the experience of AAF during non-vocal music production may not be as disruptive as for vocal production. During vocal production, AAF leads to deviations from strongly formed perception–action associations. During non-vocal music production, however, such strong associations may not exist given the putative unreliability of such associations in daily life. Second, Howell and colleagues suggest that differences across effectors in susceptibility to disruption may be specific to manipulations of feedback contents (akin to "flicking a switch"), whereas manipulations of feedback synchrony (akin to "regulate[ing] meter") may lead to similar effects across musical keyboard production and speech (vocal production).

Thus, in comparing the effects of AAF on keyboard and singing production we were interested in two questions. First, we were interested in whether the same sequencing/timing dissociation exists across effector systems with respect to the effect of AAF. Such a result would suggest that a similar kind of representation links perception and action within both effector systems. Second, we were interested whether in the disruptive effect of AAF for each effector system differed in degree. Such a result would suggest that associations between perception and action (possibly via a shared representation) differ with respect to strength.

### 1.3. Current experiment

We addressed whether the sequencing/timing dissociation is found across effector systems, and whether disruption is similar across systems, in an experiment in which participants sang melodies or performed them on a keyboard while experiencing AAF or normal feedback. The experiment was explicitly designed to test the speculations articulated by Howell et al. with respect to the effect of AAF, but the implications clearly extend beyond the context of AAF. The issue at stake is whether perception–action planning associations are based on direct, effector-specific codes, or on more abstract codes that extend across effector systems during general sequence production tasks.

In the current work, we endeavored to model AAF conditions for use in singing task based on those used in earlier studies with keyboard production tasks. In addition, we included a constant frequency shift condition modeled on those commonly used in other studies of vocal production (e.g., Burnett, Senner, & Larson, 1997; Hain et al., 2000; Jones & Keough, 2008; Jones & Munhall, 2000; Zarate & Zatorre, 2008). For the vocal production portion, participants learned and then reproduced short, novel melodies by singing. While singing they could hear feedback that was normal, frequency shifted, shifted to a random degree at every onset, or timed in such a way as to model asynchronous or serially shifted AAF. The constant shift and random shift AAF conditions further serve to test the degree to which singing might be more sensitive to AAF than keyboard production, because these conditions yield small or negligible effects in keyboard production (Pfordresher, 2005, 2008). By contrast, frequency shifted feedback does influence vocal production, though it is not clear whether these effects are “disruptive” to performance in the way that other AAF conditions are. For instance, effects of frequency-shifted feedback are often interpreted as “adaptation” rather than “disruption” (e.g., Hain et al., 2000).

A methodological difficulty in manipulating AAF for vocal production is the introduction of manipulations that only influence feedback contents while maintaining synchrony between actions and sound. In comparison with keyboard production, MIDI software (e.g., Finney, 2001) allows the user to present frequency-altered feedback in synchrony with the onset of a key press. Yet, in vocal production such fine controlled, temporally precise pitch manipulation is difficult to implement. As such, in order to manipulate pitch contents for vocal production trials, we utilized a simple method of pitch delay that was timed to match inter-onset intervals set by a metronome before each trial. This manipulation validly replicates the experience of serially shifted AAF insofar as participants maintain the tempo, a matter we further discuss in Section 3.

## 2. Method

### 2.1. Participants

Fifteen students from the University at Buffalo participated in the experiment. The participants' mean age was 22 years, with a range of 18 to 34 years. Eight participants were female and 7 were male. All participants reported being right handed, and no participants reported having hearing problems or absolute pitch.

Participants were sampled without regard for musical training from a population diverse with respect to socio-economic status. As a result, most participants had minimal to no musical training. Eight participants reported no training or musical experience on any instrument, and only minimal experience singing (one of these participants reported 3 months of school chorus). Four other participants were considered musicians. Of these, one participant (the 34-year-old) reported 11 years of training that included 5 years of vocal training (20 years experience singing) and one year of piano training (5 years experience), as well as 4 other instruments. Of the remaining three musicians, two reported 5 years of training on the piano but no singing, and the final musician reported no formal training but reported performing 40 h a week currently (20 h of singing and 10 h of piano). The remaining 3 participants reported middling levels of musical training (1 or 2 years of formal training) on various instruments but no formal training on the piano or in voice.

### 2.2. Materials

Two 8-note melodic sequences were used; all participants performed both melodies, one by singing and the other on a keyboard (the assignment of melody to task domain was counterbalanced to counteract any differences in difficulty), in different halves of the experiment. Melodies comprised 5 pitch classes (C to G in the diatonic C-major scale); one melody began on C and featured a scalar contour [C E F G F E D E] whereas the other began on G and featured an alternating contour [G E F D C E D F]. Constraints on the number of pitch classes were included so that performers would not have to change hand positions when performing on a keyboard; furthermore the use of pitches within a 7-semitone range is within the spontaneous vocal range of most untrained singers (Pfordresher & Brown, 2007; Welch, 1979). Within these constraints, melodies were designed to be maximally distinct.

We presented melodies differently for blocks of trials that involved singing or keyboard production based on pilot studies, in order to optimize learning within each effector system. For singing trials, melodies were presented aurally using a synthesized voice with pitches and formant structures designed to mimic the female or male voice (matched to gender of the participant). Participants learned melodies by imitating these synthesized recordings (see Section 2.5, for further details). We felt that imitation was an easier, more natural vocal learning procedure than the presentation of music notation.

We originally intended to have participants learn keyboard melodies also by imitation. However, pilot studies suggested that participants found learning keyboard melodies in this way more difficult than learning sung melodies through imitation (and more difficult than learning keyboard melodies via notation). This difference makes intuitive sense considering that the spatial mapping of actions on a keyboard may be better suited to notational learning than imitative learning. Because we were more interested in equating ease of learning than we were in equating the modality through which learning occurs, we decided to have participants learn keyboard melodies through simplified notation instead (first described in Pfordresher, 2005). Melodies were presented as rows of numbers beneath images of the right hand with the relevant finger highlighted (1 = thumb of right hand), to allow for participation from individuals without musical training. Thus, the first melody given above was displayed as the row [1 2 3 5 4 3 2 3] and the second melody was displayed as [5 3 4 2 1 3 2 4]. On the keyboard, numbers 1–5 were arranged in a row above the corresponding piano keys, with arrows pointing to the requisite piano key.

### 2.3. Conditions

Participants experienced one of 5 different auditory feedback conditions on a trial. Auditory feedback was presented over headphones. During singing tasks, auditory feedback was presented along with pink noise used to mask air-conducted feedback (masking noise was presented at approximately 74 dB SPL, A-weighting, against a background intensity level of approximately 37 dB SPL; the intensity of target stimuli was approximately 82 dB SPL). Auditory feedback conditions included a normal feedback control and 4 altered auditory feedback (AAF) conditions: asynchronous AAF, serially shifted AAF, random shifts of pitch and a constant frequency shift.

AAF conditions during keyboard trials were implemented using the software package FTAP (Finney, 2001), which manipulates auditory feedback via MIDI. During asynchronous AAF conditions the MIDI-out signal after each key press was delayed by 300 ms. During serially shifted AAF the event number (pitch) associated with each key press was stored in a buffer and output at the following key press; the only asynchronies in such cases result from normal MIDI lags which according to acoustic measurements in our lab were approximately 20 ms in addition to any delay amounts reported here (lags of 30 ms were reported in Repp & Keller, 2008), and similar transmission lags were measured for the apparatus used for singing trials. Such lags are not noticeable and do not lead to disruption of production, though they may have subtle effects on timing (Madison & Merker, 2004). During random pitch trials each key press led to the random selection of a MIDI note from a range of ±5 semitones around the current performed pitch, which leads to an atonal feedback melody in a pitch range characteristic of most melodies. For fixed-shift trials, participants heard feedback events that constituted a transposition of the intended melody upwards by 3 semitones.

For singing tasks, AAF manipulations were based on the auditory signal using the software package Cubase (Steinberg Media Technologies, Hamburg, Germany). Participants heard only altered feedback (with masking noise) over headphones and the loudness level was set to minimize access to air-conducted normal feedback without causing discomfort. Asynchronous and serially shifted AAF conditions were simulated by using a delay time equal to the inter-onset interval of the notes in the sequence. Specifically, the 300 ms Delay condition was designed to model an asynchronous delay, whereas the 600 ms Delay condition was designed to model the serial shift (given the prescribed IOI of 600 ms). Two further AAF conditions were used to test whether the restrictions on disruptive AAF conditions found for keyboard tasks also hold for singing. For the random pitch condition, feedback pitch was shifted to random levels (within ±3 semitones) every 600 ms.[3] The fixed-shift condition, as for keyboard production, involved a constant change in F0 equivalent to 3 semitones up.

The use of different methods for manipulating feedback across effector systems was done to maximize the validity of AAF manipulations with respect to the temporal coordination of perception and action. As mentioned before, we know of no way to sequentially manipulate AAF for singing with as much temporal precision (e.g., using MIDI) as has typically been done for keyboard production. AAF manipulations for singing in the current study thus constitute the closest approximation we could think of to the serial shift used in keyboard production. Of course, it would have been possible to manipulate serially shifted AAF during keyboard tasks in the same way as we did for singing trials, which would have sacrificed temporal control in keyboard tasks but would have maintained parity across tasks with respect to the manner in which AAF manipulations were carried out. In fact, we did just this in an aforementioned pilot study (part of which was reported by Pfordresher & Varco, 2010). Unfortunately, timing variability in keyboard performance was found to be unacceptably high (considerably higher than in singing) in order to maintain the internal validity of AAF manipulations. That is, if we were to manipulate AAF in keyboard tasks using delays of the auditory signal, serially shifted feedback would be noticeably asynchronous with key presses most of the time. Based on these facts we chose to use manipulations that best controlled coordination of actions with feedback across tasks by varying the manner in which feedback was manipulated.

[3] The difference across tasks with respect to the extent of pitch deviations was an experimental oversight. However, a pilot study that included exactly the same randomization manipulation across both tasks yielded the same kind of results we report here.

### 2.4. Apparatus

#### 2.4.1. Singing tasks

Singing tasks took place inside of a sound attenuated room (Whisper Room Inc., SE 2000 Series, Morristown, TN). The stimuli were generated using Yamaha's Vocaloid Leon software package (Zero-G Limited, Okehampton, UK) and presented over Sennheiser HD 280 Pro headphones. Participants were recorded using a Shure PG58 microphone into a Lexicon Omega recording interface. Cubase was used to present stimuli and feedback alterations, and to record imitations. Two VST plug-ins, Cakewalk Delay and de la Mancha pitchfork, were used within Cubase to alter the audio in real-time.

#### 2.4.2. Keyboard tasks

For keyboard tasks, participants used an M-AUDIO Keystation 49e unweighted piano keyboard positioned at a comfortable height. Notation for stimuli was represented as a number row (e.g., "1 2 3 5 4 3 2 3"), where 1 indicates the thumb and 5 indicates the pinky. On the keyboard, numbers 1–5 were arranged in a row above the corresponding piano keys, with arrows pointing to the requisite piano key. The software program FTAP (Finney, 2001) was used to manipulate auditory feedback, to acquire MIDI data, and to control a Roland RD-700 digital piano that produced auditory output. Participants heard auditory feedback and metronome pulses over Sony MDR-7500 professional headphones at a comfortable listening level. The piano timbre originated from Program 1 (Standard Concert Piano 1), and the metronome timbre originated from Program 126 (standard set, MIDI Key 56 = cowbell) of the RD-700.

### 2.5. Procedure

Singing and keyboard production trials were conducted in separate halves of the experiment, and the ordering of each half was counterbalanced across participants. Before singing trials, participants engaged in warm-up exercises that included singing "Happy Birthday," generating vocal sweeps (continuous oscillations of F0 between the highest and lowest comfortable pitch), and sustaining pitches representative of one's "comfort pitch." These tasks were used to avoid vocal artifacts such as creaky voice and did not involve any learning of the sung sequences. The warm-up phase was followed by a learning phase in which participants listened to, and then sang along with, a recording of a synthesized voice that repeated the sequence a total of 6 times. The participants could repeat the learning phase as much as needed until the participant and experimenter were both convinced that the sequence was memorized; all participants had it memorized by the end of the initial learning phase and participants typically found sequences easy to learn. A single practice trial and 10 experimental trials followed this learning phase. On each trial, participants heard the learned sequence one time along with a metronome set at 100 BPM (600 ms IOIs). Just after listening to the sequence, participants began vocal production. The metronome continued for four beats after the sequence while the participants imitated the sequence repeatedly for 37 s (chosen to allow around 7–8 times through the sequence). After one repetition of the sequence, one of the experimental feedback conditions (normal or AAF) took place for the remainder of the trial. During the practice trial participants heard a delay of 900 ms (a combined asynchrony and serial shift), which was not experienced in the rest of the experiment. The ten experimental trials comprised two random orders of the 5 feedback conditions presented in succession so that all feedback conditions were experienced before any one condition was repeated. Participants were instructed to use the syllable "la" when singing each note.

Keyboard trials, which could occur before or after singing trials (depending on counterbalancing order), were conducted in the same way as singing trials except that they did not include a warm-up phase. Instead, participants were presented with the simplified notation format and the experimenter made sure participants understood how to perform the notated melody (all participants did). The participants then memorized the melody in view of the notation. Participants were instructed to perform melodies at a moderate tempo, with rhythmic regularity, and using legato (connected) articulation. Following learning and memorization, the notation was removed and participants performed from memory for the remainder of the keyboard trials, which included a practice trial and 10 experimental trials that conformed to a different random order than was used for singing trials.

### 2.6. Data processing

Keyboard events were collected from the MIDI data stream through FTAP, which encodes onset times for key presses and MIDI note numbers.

Sung events were analyzed in the following way. Event onsets were demarcated via annotations with Praat (Boersma & Weenik, 2008). In determining onsets, we used information about the pitch trace, amplitude, and spectrum, all of which fluctuated with syllable boundary. These demarcations of beginnings were used to determine inter-onset intervals (IOIs) in production. Event pitches were estimated within these boundaries by finding the median F0 within the middle 50% of all pitch samples between boundaries. This procedure was effective at reducing the impact of outliers in estimated F0 (from artifacts in pitch tracking or aberrant vocal production) and at limiting the influence of vocal "scoops" which could occur near the beginning and end of sung events. All extracted estimates of pitch were conducted with in-house Matlab scripts (Mathworks, Natick, MA) and were visually inspected for accuracy.

Measures of timing for both tasks were based on inter-onset intervals (IOIs). For measures of accuracy, we focused on the produced melodic contour, the pattern of upward or downward changes between successive pitches, irrespective of interval size. Our use of melodic contour was based on the fact that singers frequently mistune notes and/or distort interval size in production, but rarely make errors of melodic contour, when singing with normal feedback (Pfordresher & Brown, 2007). We were interested in disruption specific to AAF here, rather than error patterns that may simply reflect individual differences in intonation production. Thus, each melodic interval produced by voice or on the keyboard was coded as $+1$ if ascending and $-1$ if descending, and compared to a similar ordinal coding of melodic intervals in the target stimulus. The Unix program "diff" was used to assess the frequency of addition, deletion, or substitution errors. The diff program operates under similar principles to other programs used to assess ordering errors in music (Large, 1993; Palmer & van de Sande, 1993, 1995) but does not make any assumptions regarding error sources, event timing or key. We removed any errors at the end of the trial that were byproducts of earlier errors in the trial (e.g., if a single note event is added in the middle of a trial, then it might appear as though the final event was deleted).

### 2.7. Design and analysis

The complete experimental design included the within participants factors task (keyboard, singing) and feedback condition (normal, asynchronous AAF, serially shifted AAF, fixed shift AAF, random AAF). Between participants conditions were used for counterbalancing and included the factors block order (singing first, or keyboard first), assignment of melody to task domain (2 levels), and random order of trials (2 orders). None of the counterbalancing variables yielded significant effects. Thus we analyzed dependent measures using 2-way within-participants analyses of variance (ANOVA). Our primary measures of disruption were mean IOI (timing) and the proportion of contour errors in a trial (accuracy), as in previous studies on the effects of AAF.

In addition to omnibus ANOVAs we incorporated complex planned contrasts within each task (Keppel & Wickens, 2004, pp. 76–83) based on the sequencing/timing dissociation hypothesis. Specifically, when analyzing timing we assessed the hypothesis that asynchronous feedback elevated the dependent measure relative to all other feedback conditions within each task by separating data across keyboard and singing trials and applying the coefficients $[(-1*\text{normal})+(+4*\text{asynchronous AAF})+(-1*\text{serially shifted AAF})+(-1*\text{fixed-shift AAF})+(-1*\text{random AAF})]$ to the mean of each condition. When analyzing errors we adopted a similar procedure but instead assigned a coefficient of $+4$ to serially shifted AAF and $-1$ to the other conditions. Because complex contrasts were designed to test effects within a task domain, the error term was based only on variability within that task domain; this was computed by running an ANOVA only on one task domain (singing or keyboard production), which was conducted strictly for the purpose of computing the complex contrast.

## 3. Results

### 3.1. Disruptive effects of AAF

Fig. 2 displays the effect of task and feedback condition on timing (2A) and accuracy (2B). The ANOVA on mean IOI (Fig. 2A) yielded a significant main effect of task, $F(1, 14)=5.76$, $p<.05$, and a significant main effect of feedback condition, $F(4, 56)=20.45$, $p<.01$, but no significant interaction ($p=.97$). Complex contrasts based on the hypothesis that asynchronous feedback selectively disrupts timing were significant for both keyboard performance, $F(1, 56)=33.55$, $p<.01$, and singing, $F(1, 56)=57.76$, $p<.01$. Participants performed the keyboard at a faster rate than they sang, and performed more slowly with asynchronous feedback than any other condition in both tasks. Percents above means in Fig. 2A are used for illustrative purposes, to show the disruptiveness of each AAF condition relative to performance with normal feedback. These figures suggest highly similar effects of asynchronous feedback in each condition. During piano trials, participants performed faster than the intended IOI; however, this difference did not compromise the AAF manipulations used in this task.

The mean percent of contour errors in a trial are shown in Fig. 2B. The ANOVA yielded a significant main effect of task, $F(1, 14)=28.45$, $p<.01$, a main effect of feedback, $F(4, 56)=17.08$, $p<.01$ and a significant task × feedback interaction, $F(4, 56)=12.24$, $p<.01$. Participants made more errors overall when singing than when playing the keyboard and in addition to this the disruptive effects of serially shifted AAF was relatively greater in singing than in keyboard performance (see percents above means). Despite these differences in susceptibility to disruption across effector systems, complex contrast analyses suggested that serial shifts had similar qualitative effects within each system, in that serial shifts selectively disrupted accuracy for both keyboard performance, $F(1, 56)=7.52$, $p<.01$, and singing, $F(1, 56)=57.81$, $p<.01$.

### 3.2. Error types

The fact that similar patterns of errors were found across singing and keyboard tasks does not necessarily mean that the same types of errors occurred in each task. As such, Fig. 3 shows mean numbers of errors across different error types, including additions, deletions and substitutions. Most errors were either additions (3A) or deletions (3B) and so it is not surprising that patterns of these errors closely resembled overall error rates. However, substitution errors for singing trials differed from the general pattern of disruption in that this measure did not show the same kind of disruptive effect for serial shifts as did the other measures. By contrast, substitution errors for keyboard trials, though infrequent, mirrored the pattern seen for overall errors.

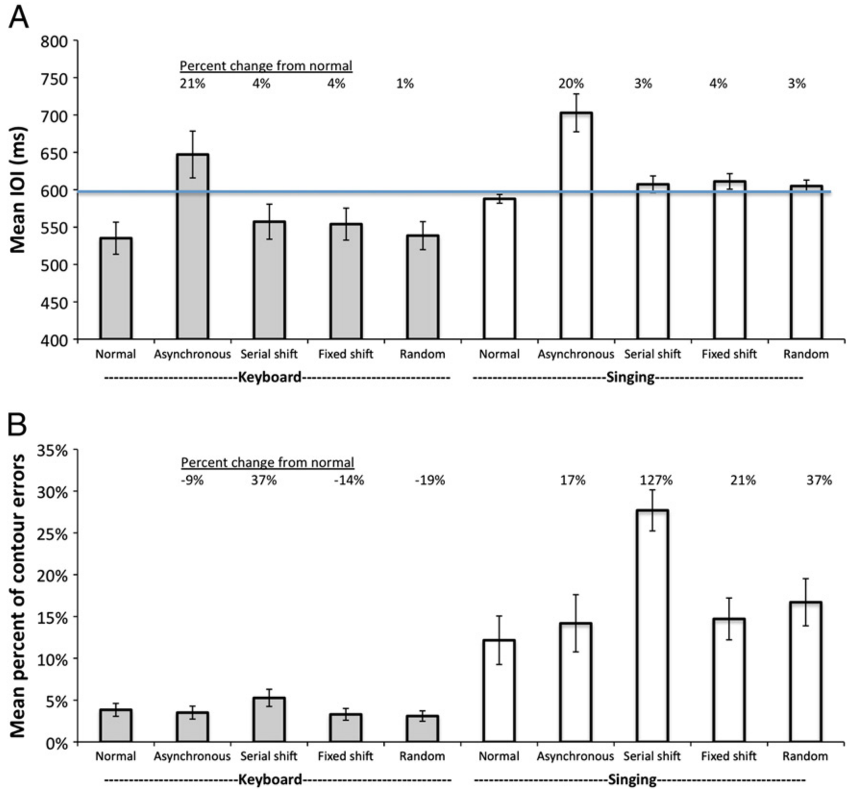


**Fig. 2.** Effects of auditory feedback condition and task (keyboard versus singing) on timing (A) and the accuracy of sequencing (B). Bars show means across participants and trials, with error bars representing the standard error of the mean. Numerical values above the means for AAF condition represent normalized disruption measures. The horizontal line in panel A represents the prescribed (ideal) performance tempo.

Thus far our analyses of errors have focused on deviations from melodic contour. Such deviations in relative pitch are most central to the research reported here; nevertheless it is also of interest to determine whether AAF manipulations influence participants' tendency to sing in tune overall (cf. Pfordresher & Brown, 2007). We addressed this issue by taking the absolute value of the difference between the mean of all produced pitches on a trial and the mean pitch of the target stimulus (in cents). We computed pitch in this way, rather than on a note-by-note basis, in order to examine overall mistuning independently of errors that occur on individual notes. Means for these absolute differences across participants are shown in Fig. 4. We only examine performance in singing trials because no tendency to mistune overall was found in keyboard trials, where participants produced sequences using 5 fixed pitches. As can be seen, the pattern of results differs considerably from errors based on relative pitch. All feedback conditions were equivalent with the exception of the fixed pitch shift condition. The ANOVA on these means was not significant ($p > .20$). However, we performed a complex contrast based on the prediction that mistuning would be greatest for the fixed pitch shift condition (which follows from previous research, e.g. Jones & Keough, 2008), and this contrast was significant, $F(1, 56) = 4.84$, $p < .05$. Thus, we can conclude that serially shifted AAF disrupts the sequencing of events, but not necessarily one's tendency to sing in tune, whereas fixed pitch shifts cause one to sing out of tune but does not disrupt sequencing.

### 3.3. Effects of musical training

The majority of our participants had minimal or no musical training. However, as mentioned earlier 4 participants should be considered musicians. Moreover, it is of theoretical interest to see if the effects of AAF documented here for both effector systems are at all modulated by musical training. In order to address this issue as cleanly as possible we analyzed separately data from 8 participants who had no musical training at all. Their data are shown in Fig. 5 (white bars). For illustrative purposes we also plot the median and individual means from the 4 participants who could be considered musicians; we did not analyze the musicians parametrically given the sample size. As can be seen the basic results found for the 8 untrained participants greatly resemble the data from all 15 participants, and are for the most part similar to the data for the 4 musicians. Two exceptions are that the non-musicians made more errors overall when singing (though not when playing the keyboard), and that timing of keyboard performance among musicians was slightly slower for the fixed and random shifts of pitch. As a whole, however, these data argue against the idea that perception–action interference effects from AAF are based on acquired associations during musical training.

Within non-musicians, statistical effects were similar to those found for the entire sample. The ANOVA on timing yielded only a significant main effect of feedback, $F(4, 28) = 16.88$, $p < .01$, and complex contrasts were significant for keyboard trials, $F(1, 28) = 33.55$, $p < .01$, and singing trials, $F(1, 28) = 57.76$, $p < .01$. The ANOVA on error rates yielded a significant main effect of task, $F(1, 7) = 24.62$, $p < .01$, a significant main effect of feedback, $F(4, 28) = 7.09$, $p < .01$, and a task × feedback interaction, $F(4, 28) = 4.54$, $p < .01$. As can be seen in Fig. 5B, the interaction supports the same pattern of results as found for the entire sample, with the effect of serially shifted AAF being greater within singing trials than within keyboard trials. Nevertheless, as with the entire sample, complex contrast analyses revealed disruptive effects of serially shifted feedback both for keyboard trials, $F(1, 28) = 15.73$, $p < .01$, and for singing trials, $F(1, 28) = 37.55$, $p < .01$.

### 3.4. Internal validity of AAF manipulations

Finally, we report the analyses designed to test the validity of serially shifted AAF manipulations within singing trials. As mentioned before,

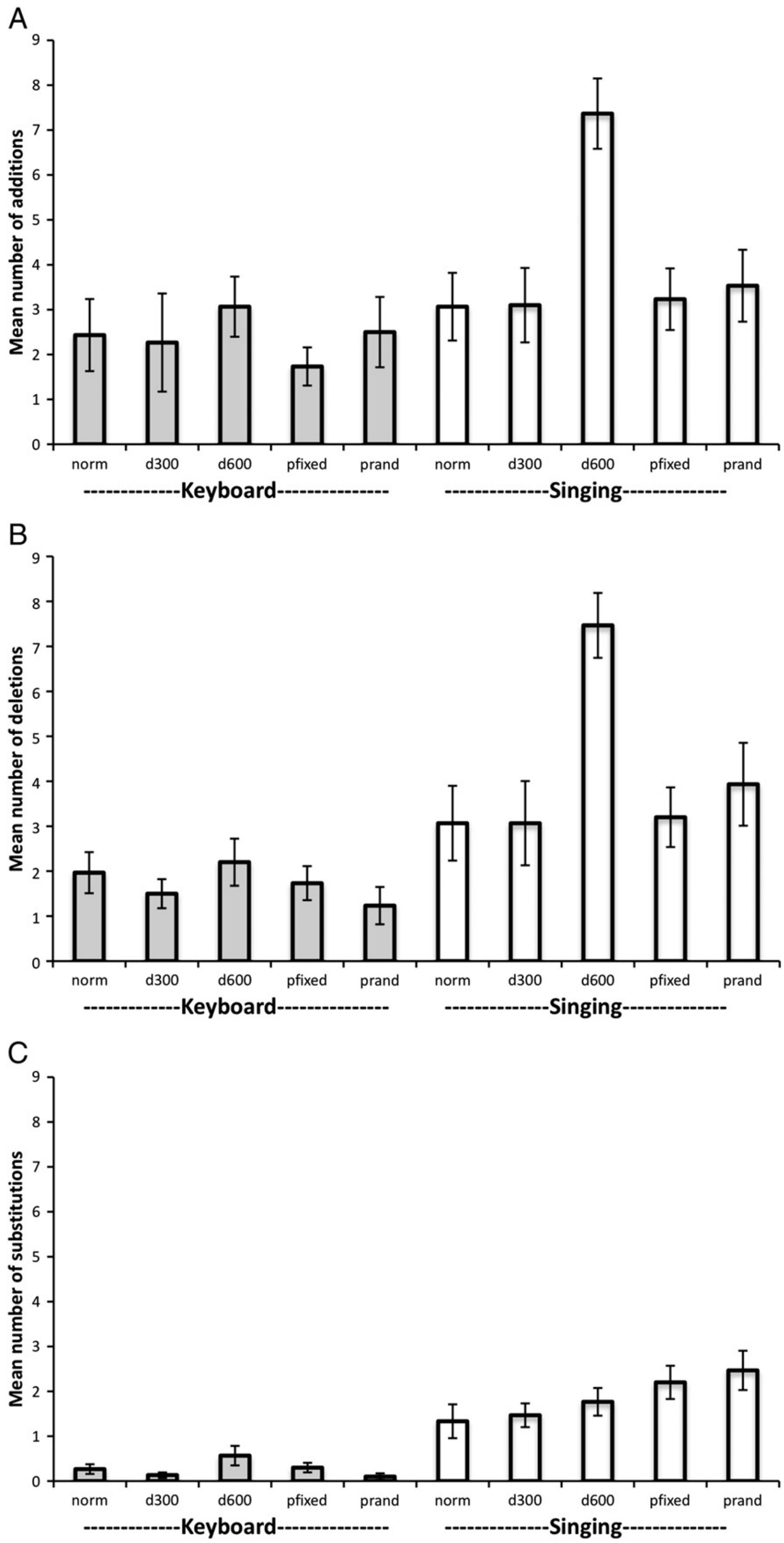


**Fig. 3.** Effects of auditory feedback condition and task on different error types, including the number of addition errors (A), the number of deletions (B), and the number of event substitutions (C). Bars show means across participants and trials, with error bars representing one standard error of the mean.

the validity of these manipulations depends on how well participants maintained the prescribed tempo of 600 ms per IOI. Deviations from this rate lead to asynchronies; if these are large enough then one could say that disruption may reflect the combined effect of asynchronous and serially shifted AAF. Past research suggests that asynchronies between actions and an accompanying auditory pattern (a metronome,

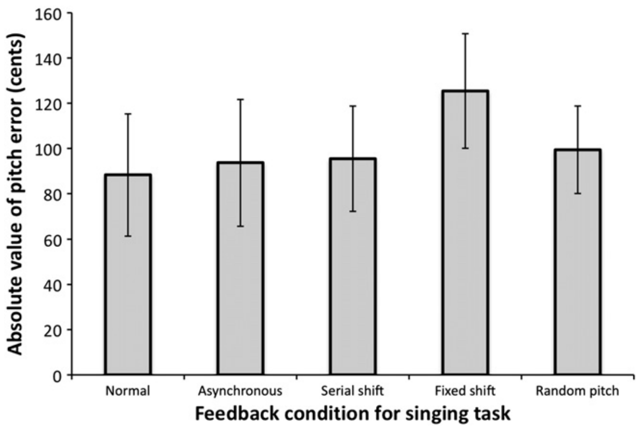


**Fig. 4.** Effects of feedback condition on the absolute difference between mean sung F0 and the mean F0 in target sequences (a measure of mistuning) within singing tasks. Bars show means across participants and trials, with error bars representing 1 standard error of the mean.

or auditory feedback) may not be detectable even up to 100 ms (Jäncke, 1989; cited in Aschersleben & Prinz, 1997). In order to address this issue, we categorized each produced IOI during singing trials with the (simulated) serial shift condition into 10-ms bins and categorizing each bin according to its absolute difference from the 600 ms target IOI. Fig. 6 shows the resulting histogram, ranging from 0 to 300 ms. The lower rate for the 0-ms bin reflects the fact that all other bins reflect the sum of IOIs both shorter and longer than 600 ms. Importantly, the majority of IOIs (73%) fall within 100 ms of the target IOI. Another, more conservative, criterion is suggested by an asynchrony detection task reported by Repp (2000, Experiment 3) who found that asynchronies up to 10% of the target IOI were not detectable by most participants. In all, 60% of IOIs fell within this boundary (an absolute difference of 60 ms). Given these figures one can assume that most asynchronies within the (simulated) serial shift condition for singing trials were not detectable. Moreover, even those that could be detectable would be accompanied by alterations to feedback contents and would be considerably less asynchronous than the 300 ms asynchrony (50% of prescribed IOIs).

Another claim on which the current method was based has to do with timing variability across the two tasks. Recall that our motivation for using more temporally controlled manipulations in the keyboard tasks than in the singing tasks was that timing variability in production is greater in the former than the latter (see Section 2.3). We assessed this claim by analyzing coefficients of variation across both tasks (standard deviation normalized by mean IOI). Timing variability in keyboard tasks (M CV = 84%, SE = 2%) was several orders of magnitude higher than in singing tasks (M CV = 18%, SE = 1%), and the difference, not surprisingly, was statistically significant, $t(14) = 18.38$, $p < .01$. Thus, if we had manipulated serial shifts in keyboard tasks as we had done for singing tasks we would most likely not have replicated the sequencing/timing dissociation for keyboard tasks, a speculation that was confirmed in the aforementioned pilot study.

## 4. Discussion

The results of the experiment reported here suggest that the sequencing/timing dissociation generalizes across vocal and manual effector systems during the production of musical sequences. Similarity across effector systems was particularly striking for the effect of

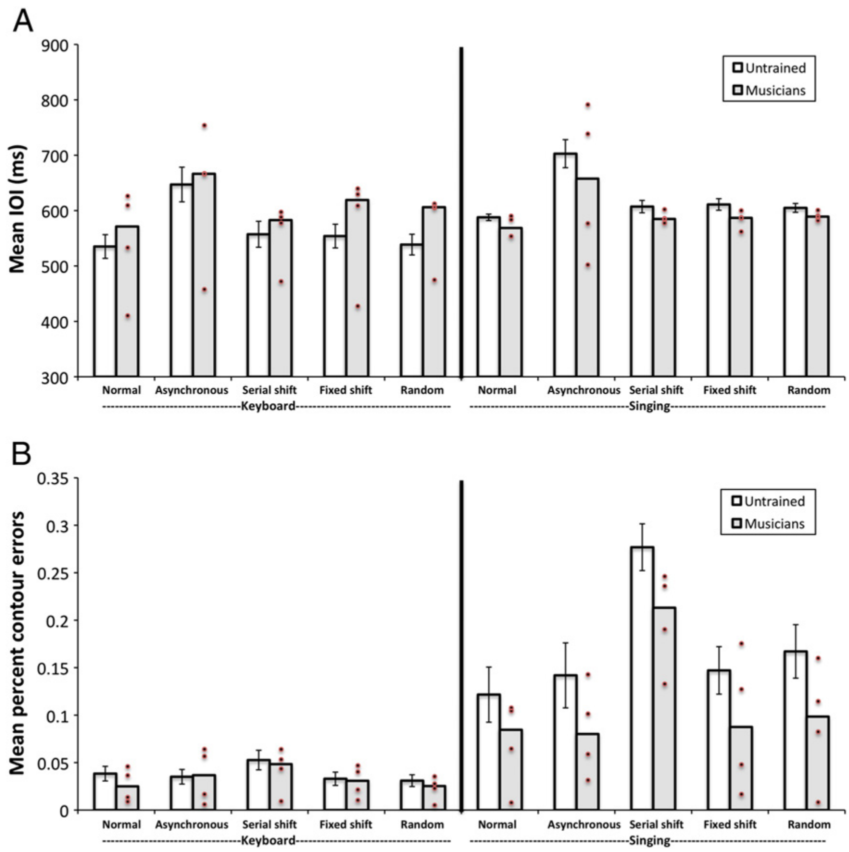


**Fig. 5.** Effects of auditory feedback condition and task on timing (A) and sequencing (B) for different groups of participants defined by levels of musical training (untrained participants versus musicians). White bars represent means (across participants and trials) among 8 musically untrained participants, with error bars representing 1 standard error of the mean. Gray bars represent medians across 4 musicians, whose individual means are shown as dots surrounding each gray bar.

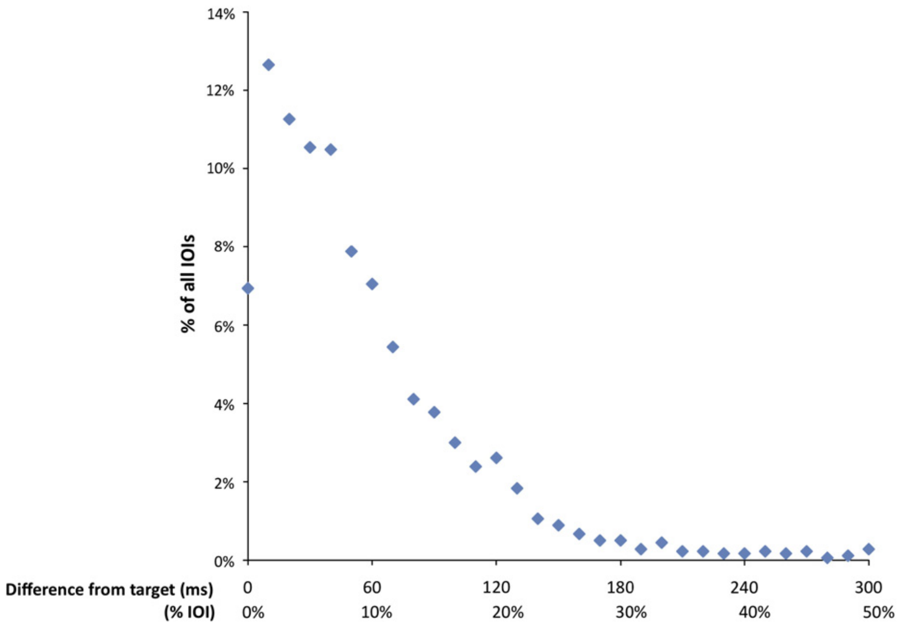


**Fig. 6.** Histogram showing the distribution of all inter-onset intervals (IOIs) within singing trials in which participants experienced serially shifted AAF (a delay of 600 ms). Each produced IOI was classified within bins of 10 ms; the figure plots centers of bins. IOIs are plotted as deviations from the target IOI of 600 ms (defined by the tempo); the abscissa shows deviations as absolute values and as percents of the target IOI. Every bin other than 0 ms reflects the sum of IOIs that are longer and shorter than the target.

asynchronous AAF on timing (Fig. 2A), which yielded almost identical effects across tasks. Indeed, this finding is in line with the earlier prediction by Howell et al. (1983) who proposed that asynchronous feedback (or any recurring sound that is asynchronous with actions) should disrupt the timing of sequence production in any effector system. Furthermore, both effector systems showed the sequencing/timing dissociation, found previously for keyboard production (Pfordresher, 2003; Pfordresher & Kulpa, 2011). Asynchronous AAF disrupted timing, but not accuracy, whereas serial shifts of pitch disrupted accuracy, but not timing. These qualitative similarities across systems are consistent with the idea that perception–action associations are formed at an abstract level, possibly based on common coding of goals associated with action and resulting perceptual events, rather than effector-specific associations. At the same time, effector systems differed dramatically in their susceptibility to disruption by serial shifts. This quantitative difference follows logically from the fact that motor gestures in the vocal system are more reliably linked to actions than are manual action sequences (cf. Howell et al., 1983).

Although the effector systems explored here differ with respect to their susceptibility to disruption by serial shifts, it is noteworthy that these systems did not differ with respect to the potentially disruptive effects of other alterations to feedback contents, namely random and constant shifts of frequency. It might be predicted, for instance, that the tighter link between actions and sounds in vocal production would lead to disruption of vocal production from random pitch shifts as well as from serial shifts of pitch. This was not the case. Thus, the systems do not appear to differ with respect to the role of feedback contents in the sequencing of movement. Moreover, it is interesting to note that the effects of constant frequency shifts of vocal feedback are not “disruptive” in the sense of leading to more errors. Although such manipulations do lead to alterations in produced pitch (and they had such effects here too), such effects amount to musical “transpositions” in singing rather than disruption in the production of pitch contour.

Our inclusion of participants who were musically untrained bears discussion. Had we sampled only pianists we could have benefitted from their greater temporal precision in production and have possibly been able to fully equate AAF manipulations across tasks. However, for theoretical reasons it was important to us that we include untrained participants, due to the aforementioned point made by Howell et al. (1983). Assuming their claim about perception–action relationships being arbitrary in musical (non-vocal) performance, an untrained individual could be assumed to have little or no sense for what kind of perception–action relationships are appropriate. Such an individual should not be disrupted simply because of hearing an inappropriate pitch if no pre-existing associations between piano keys and pitches had been formed. For instance, previous reports of associations between pitch height and single, discrete responses have found reduced effects, limited only to pitch changes greater than an octave (Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006; see also Keller & Koch, 2008). Thus, the use of non-musicians constitutes a strong test of the hypothesis given by these researchers. By contrast, were we to use trained pianists, then this group might have developed associations over time out of perception–action associations that may have initially been treated as unrelated. If we found results like the current results with a group of pianists, one could claim that the effects simply reflect learned associations for keyboards, combined with more intrinsic associations with the voice.

Although we attempted as much as possible to equate conditions across keyboard and vocal tasks, there were many respects in which these tasks differed beyond simply the motor system used. These differences included the modality used to present stimuli during learning (auditory for singing, visual for keyboard), the method used for manipulating serial shifts (based on time delays for singing, based on MIDI for the keyboard), the method for extracting produced pitch (analysis of F0 for singing, use of MIDI event codes for keyboard), and the use of warm-up trials (adopted for singing but not keyboard trials). In our view all of these differences were necessary based on the important goal of ensuring that the task demands and feedback manipulations were best suited for the motor task accomplished by our largely untrained participants. Moreover, it is difficult for us to see how these task differences could provide a successful alternative account of the current results. First, our primary result is one of similarity across task domains. We feel that it is unlikely that these similarities could result as an artifact of differences in our manipulations. If anything these differences across domains should enhance different responses to AAF. This leads to the second primary finding, which is that singing was evidently more vulnerable to the effect of serial shifts than keyboard production. Perhaps the slight asynchronies that were present in this manipulation for singing (which simulates the kind of manipulation we could execute more cleanly in keyboard production) enhanced disruption for singing tasks. Yet there is a problem with this account as well. That is, asynchronies typically lead to slowing of production compared to normal feedback (this was found for singing and keyboard tasks with the 300 ms delay); yet the (simulated) serial shift condition for singing did not slow production (see Fig. 2A). Thus, our interpretation of the results is that we found similar effects in both task domains in spite of necessary differences in the procedure we used across tasks. To us the similar results argue for the robustness of the sequencing/timing dissociation. Finally, the presence of higher error rates during normal feedback for the singing condition may be taken to suggest that this motor task was more difficult for participants, and indeed many people have difficulty singing accurately and precisely (Pfordresher, Brown, Meier, Belyk, & Liotti, 2010). However, anecdotal observations by experimenters suggested that the learning of musical sequences (which is separate from intonation) was in fact easier for singing than for keyboard tasks.

In conclusion, the current results favor the view that the disruptive effects of AAF are based on abstract, effector-independent, associations between perception and action. These connections may not rely on direct experience with the relationship between actions and sound within a particular medium. Instead, we think that AAF manipulations reflect the fact that people in general learn to expect that the physical pattern of actions they generate under normal circumstances should be correlated with the pattern of perceptual events that accompany actions. That this sensitivity is (apparently) not contingent on registering perceptual events as feedback or on the effects of learning suggests that we rely on such patterns of coordination at a very basic level.

## Acknowledgments

This research was sponsored in part by NSF grant BCS-0642592. We are grateful to Anastasiya Kobrina for help with data analysis and to Thomas Varco who arranged the singing AAF paradigm.

## References


Aschersleben, G., & Prinz, W. (1997). Delayed auditory feedback in synchronization. *Journal of Motor Behavior*, *29*, 35–46.

Effects of altered auditory feedback in music performance: The role of impulsivity, musical aptitude, and general intelligence. Unpublished Master's Thesis, University of Texas at San Antonio, San Antonio, TX.

Black, J. W. (1951). The effect of delayed side-tone upon vocal rate and intensity. *The Journal of Speech and Hearing Disorders*, *16*, 56–60.

Boersma, P., & Weenik, D. (2008). *Praat: Doing phonetics by computer (Version 5.0.25).*

Borden, G. J. (1979). An interpretation of research on feedback interruption in speech. *Brain and Language*, *7*, 307–319.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience*, *13*, 400–404.

Burnett, T. A., Senner, J. E., & Larson, C. R. (1997). Voice F0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice*, *11*, 202–211.

Cohen, A., Ivry, R. B., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 17–30.

Drost, U. C., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *The Quarterly Journal of Experimental Psychology*, *60*, 527–533.

Fairbanks, G., & Guttman, N. (1958). Effects of delayed auditory feedback upon articulation. *Journal of Speech and Hearing Research*, *1*, 333–346.

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, *15*, 153–174.

Finney, S. A. (1999). *Disruptive effects of delayed auditory feedback on motor sequencing.* Providence: Brown University.

Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, *33*, 65–72.

Finney, S. A., & Palmer, C. (2003). Auditory feedback and memory for music performance: Sound evidence for an encoding effect. *Memory & Cognition*, *31*, 51–64.

Gates, A., Bradshaw, J., & Nettleton, N. (1974). Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, *14*, 21–25.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *The Journal of Neuroscience*, *18*, 9420–9428.

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

Havlicek, L. (1968). Effects of delayed auditory feedback on musical performance. *Journal of Research in Music Education*, *16*, 308–318.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *The Behavioral and Brain Sciences*, 849–937.

Howell, P. (2001). A model of timing interference to speech control in normal and altered listening conditions applied to the treatment of stuttering. In B. Maassen, B. Hulstijn, R. Kent, H. F. M. Peters, & P. H. M. M. van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 291–294). Nijmegen: Uttgeverij Vantilt.

Howell, P. (2004). Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering. *Stammering Research*, *1*, 34–43.

Howell, P. (2007). The effects of gated speech on the fluency of speakers who stutter. *Folia Phoniatrica et Logopaedica*, *59*, 250–255.

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics*, *36*, 296–302.

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology. Human Perception and Performance*, *9*, 772–784.

Howell, P., & Sackin, S. (2002). Timing interference to speech in altered listening conditions. *Journal of the Acoustical Society of America*, *111*, 2842–2852.

Jäncke, L. (1989). Die Bedeutung dr. auditiven Rückmeldung des eigenen Sprchens für die Kontrolle der Phonationsdauer [Influence of auditory feedback of speech on the control of phonation time]. Unpublished doctoral dissertation, University of Düsseldorf, Düsseldorf, Germany.

Jones, J. A., & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research*, *190*, 279–287.

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, *108*.

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *The Quarterly Journal of Experimental Psychology*, *61*, 275–291.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Krampe, R. T., Mayr, U., & Kliegl, R. (2005). Timing, sequencing, and executive control in repetitive movement production. *Journal of Experimental Psychology. Human Perception and Performance*, *31*, 379–397.

Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, & Computers*, *25*, 238–241.

Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, *22*, 824–826.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer-Verlag.

MacKay, D. G., & Bowman, R. W. J. (1969). On producing the meaning in sentences. *The American Journal of Psychology*, *82*, 23–39.

Madison, G., & Merker, B. (2004). Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neuroscience Letters*, *370*, 69–73.

Mürbe, D., Friedemann, P., Hofmann, G., & Sundberg, J. (2002). Significance of auditory and kinesthetic feedback to singers' pitch control. *Journal of Voice*, *16*, 44–51.

Mürbe, D., Friedemann, P., Hofmann, G., & Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback — A longitudinal study of singers' pitch control. *Journal of Voice*, *18*, 236–241.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science*, *11*, 63–68.

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 457–470.

Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology. Human Perception and Performance*, *21*, 947–962.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology. Human Perception and Performance*, *29*, 949–964.

Pfordresher, P. Q. (2005). Auditory feedback in music performance: The role of melodic structure and musical skill. *Journal of Experimental Psychology. Human Perception and Performance*, *31*, 1331–1345.

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, *2*, 183–198.

Pfordresher, P. Q. (2008). Auditory feedback in music performance: The role of transition-based similarity. *Journal of Experimental Psychology. Human Perception and Performance*, *34*, 708–725.

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "tone deafness". *Music Perception*, *25*, 95–115.

Pfordresher, P. Q., Brown, S., Meier, K., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *Journal of the Acoustical Society of America*, *128*, 2182–2190.

Pfordresher, P. Q., Cox, J. L., Andrews, M., Mantell, J. T., Brown, S., & Zivadinov, R. (2010). *Neural stratification of sequencing and timing ?* Journal of Cognitive Neuroscience, Supplement: An fMRI study 190.

Pfordresher, P. Q., & Kulpa, J. D. (2011). The dynamics of disruption from altered auditory feedback. *Journal of Experimental Psychology. Human Perception and Performance*, *37*, 949–967.

Pfordresher, P. Q., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research/Psychologische Forschung*, *66*, 71–79.

Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics*, *68*, 362–376.
Pfordresher, P. Q., & Varco, T. (2010). Altered auditory feedback effects on keyboard and singing performance. *Proceedings of the 11th international conference on music perception and cognition* (pp. 474–477). Seattle, WA: University of Washington.
Prinz, W., Aschersleben, G., & Koch, I. (2009). Cognition and action. In E. Morsella, J. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 35–71). Oxford: Oxford University Press.
Repp, B. H. (1999). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, *16*, 409–438.
Repp, B. H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research*, *63*, 106–128.
Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science*, *27*, 423–456.
Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology. Human Perception and Performance*, *9*, 86–102.
Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, *99*, 113–129.
Ward, W. D., & Burns, E. M. (1978). Singing without auditory feedback. *Journal of Research in Singing*, *1*, 24–44.
Welch, G. F. (1979). Vocal range and the poor-pitch singer. *Psychology of Music*, *7*, 13–31.
Yamaguchi, M., & Proctor, R. W. (2009). Transfer of learning in choice reaction times: Contributions of specific and general components of manual responses. *Acta Psychologica*, *130*, 1–10.
Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, *60*, 213–232.
Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*, *40*, 1871–1887.