The effects of delayed auditory and visual feedback on speech production

Jennifer Chesters,
Department of Experimental Psychology, South Parks Road, Oxford University, Oxford, OX1 3UD, UK

Ladan Baghai-Ravary, and
Oxford University Phonetics Lab, 41 Wellington Square, OX1 2JF, UK

Riikka Möttönen
Department of Experimental Psychology, South Parks Road, Oxford University, Oxford, OX1 3UD, UK

Abstract

Monitoring the sensory consequences of articulatory movements supports speaking. For example, delaying auditory feedback of a speakers' voice disrupts speech production. Also, there is evidence that this disruption may be decreased by immediate visual feedback, i.e., seeing one's own articulatory movements. It is, however, unknown whether delayed visual feedback affects speech production in fluent speakers. Here, the effects of delayed auditory and visual feedback on speech fluency (i.e., speech rate and errors), vocal control (i.e., intensity and pitch) and speech rhythm were investigated. Participants received delayed (by 200 ms) or immediate auditory feedback, whilst repeating sentences. Moreover, they received either no visual feedback, immediate visual feedback or delayed visual feedback (by 200, 400 and 600 ms). Delayed auditory feedback affected fluency, vocal control and rhythm. Immediate visual feedback had no effect on any of the speech measures when it was combined with delayed auditory feedback. Delayed visual feedback did, however, affect speech fluency when it was combined with delayed auditory feedback. In sum, the findings show that delayed auditory feedback disrupts fluency, vocal control and rhythm and that delayed visual feedback can strengthen the disruptive effect of delayed auditory feedback on fluency.

I. INTRODUCTION

Fluent speech production is a highly complex skill requiring the precise control and co-ordination of articulatory movements, voicing and respiration. Perturbation of sensory information, particularly auditory feedback of the speech signal (Lee 1950; Lee 1950; Siegel and Pick 1974; Elman 1981; Howell and Archer 1984), has been shown to disrupt characteristics of the speech produced, such as fluency, amplitude or pitch. Studies measuring the acoustic changes in speech output in response to online auditory perturbation

(e.g. shifts in formant, or fundamental frequencies), show that the adjustments made in production compensate for the perturbation (Houde and Jordan 1998; Jones and Munhall 2000; Jones and Munhall 2005; Shiller et al. 2009). The effects of these perturbations indicate that monitoring auditory feedback has a role in maintaining accurate speech production.

One way to perturb auditory feedback is to delay speaker's voice by a few hundred milliseconds. Lee (1950a, 1950b), one of the first researchers to study the effect of delayed auditory feedback (DAF) on speech production, described the speech disruptions under DAF as "stutter-like" and noted repetitions of syllables, pauses, and increases in pitch or intensity. Since then, the disruptive effects of DAF on speech have been replicated in numerous studies (e.g. Fairbanks and Guttman 1958; Yates 1963). Stuart and colleagues (Stuart et al. 2002) compared the effects on speech of DAF at various delay times (25, 50 and 200 ms). They found that a 200 ms auditory delay caused greatest disruption, indexed by increased speech errors and decreased speech rate.

Lee's early description of DAF speech as "stutter-like" is interesting, particularly as DAF can actually have the opposite effect on people who stutter, i.e., it can increase fluency (Kalinowski and Stuart 1996). In fact, there are some notable differences between speech under DAF and stuttering. For example, people who stutter tend to experience disfluency more on consonant segments (Howell et al. 1988) in word- or sentence-initial positions (Wingate 2002). In contrast, Howell (2004) noted, in his review paper "considering first the effects of DAF on fluent speakers, the most notable effect is lengthening of medial vowels", which we assume to mean vowels in the syllable nucleus. We are, however, not aware of any studies that have quantified the effects of DAF on the production of vowels and consonants. In addition, the timescales differ at which DAF modulates the speech of people who stutter compared to fluent speakers. Whereas the disruptive effects of DAF on normally fluent speakers is greatest with delay times around 200 ms, people who stutter can experience increased fluency when auditory feedback is delayed by a much shorter interval (Kalinowski and Stuart 1996).

DAF studies provide evidence that auditory feedback from the on-going speech stream is used to maintain fluent speech production. A comprehensive description of the mechanisms involved in speech production is given in models such as the Directions into Velocities of Articulators (Guenther et al. 2006; Tourville and Guenther 2011). The DIVA model provides a computational description of sensori-motor control of speech, and maps the components onto corresponding brain regions. A speech sound map, which plans speech movements, also projects to sensory target maps. These target maps act as forward models of the expected sensory outcomes of speech, and interact with state maps (which encode current state of the sensory systems), and error maps (representing the difference between state and target) to produce an error signal when the planned articulation does not match the target. According to the model, separate sets of these maps exist, for auditory and somatosensory modalities. The error signals from each modality are sent to a single feedback control map, which converts this information into a corrective motor command used to maintain accurate production. We are not aware of any studies that have simulated the effects of DAF on speech production using the DIVA model. The speech production

system as described in DIVA must be robust to some small levels of auditory delay, as there is intrinsic latency with the auditory feedback loop. A delay of 200 ms, however, would plausibly result in a detection of a mismatch between speech output and auditory feedback, and generation of large error signals. It would not be possible to use these error signals to make corrections to decrease the sensori-motor asynchrony. Instead, the error signals could act to disrupt motor control and reduce the fluency and naturalness of ongoing speech. Civier and colleagues (2010) investigated how stuttering may be accounted for within DIVA. They proposed an additional monitoring subsystem, that detects and repairs large errors for which feedback loops are unable to compensate. The monitoring subsystem is able to "reset' the speech motor system, creating a momentary interruption in the flow of speech. It may be the case that similar "resets" occur under DAF.

There is much evidence that visual information affects speech perception. For example, speech can be understood to some extent by viewing the speaker's articulatory movements alone, i.e. by speechreading (Summerfield 1992). Visual cues can enhance auditory speech perception, especially in noise (Sumby and Pollack 1954) or modify an auditory percept, as in the McGurk illusion (McGurk and MacDonald 1976). This illusion, where conflicting visual cues alter the auditory perception of a syllable, demonstrates that auditory and visual information are integrated during speech perception. Manipulation of the temporal asynchrony between the auditory and visual tokens in a McGurk paradigm suggests that the audio-visual integration window is around 200 ms in duration, and skewed toward an auditory lag (van Wassenhove et al. 2007).

Visual feedback about our own articulatory movements, unlike auditory and somatosensory feedback, is not usually available to us when we speak. However, it has been shown to aid speech production. For example, visual feedback of the speaker's face can be utilized to reduce stuttering. This effect has been demonstrated under both immediate visual feedback (IVF) and delayed visual feedback (DVF) of up to 400 ms (Snyder et al. 2009; Hudock et al. 2010). Immediate visual feedback has been recognized as a useful way to increase phonemic accuracy in acquired apraxia of speech for some years (Rosenbek et al. 1973). It has also been demonstrated that an audiovisual speech model can 'entrain' fluent speech in non-fluent aphasic patients (Fridriksson et al. 2012). In addition, therapy for childhood apraxia of speech may also use visual feedback via a mirror (Williams 2009). Little is known, however, about how visual feedback affects speech production in fluent speakers. Two previous studies have investigated how IVF may influence DAF effects on speech. Tye-Murray (1986) measured sentence durations when participants spoke under DAF whilst looking into a mirror, but found no change compared to when no visual feedback was given. However, Jones and Striemer (2007) found that IVF could affect the number of speech errors produced. When participants were divided into two groups based on their susceptibility to DAF alone, the "low disruption" group produced fewer errors when IVF was available (Jones and Striemer 2007). It was speculated that this group may have been more able to utilize alternative sources of feedback. Thus, the reduced susceptibility to DAF may be due to using somatosensory feedback, in addition to visual feedback when it was available. So, evidence for effects of IVF on the speech production of fluent speakers is mixed, and the effects of DVF on this population have not been investigated to date. The DIVA model does not include visual speech components. However, visual speech information is used during

language acquisition (Kuhl and Meltzoff 1982; Weikum et al. 2007) and evidence from patient groups confirms that visual feedback can also benefit an impaired mature speech production system. It is possible that the speech production system acquires maps and mappings regarding visual speech during development. This mechanism may then be exploited in adulthood to support impaired speech production, e.g. in stuttering and acquired speech disorders. Perturbation of the visual feedback of fluent speakers, by introducing delays, could provide important information about how the speech production system utilizes visual information.

The first aim of this study was to investigate whether DVF (combined with immediate AF and DAF) would influence speech fluency. The auditory state map described by the DIVA model is located in the superior temporal regions (Guenther et al. 2006; Tourville and Guenther 2011). Visual speech information has also access to the these auditory regions (Calvert et al. 1997; Mottonen et al. 2002), providing a route by which visual feedback may be integrated with auditory feedback. Alternatively, there could be a separate target and error maps for visual feedback. We hypothesized that if visual feedback is integrated with auditory feedback, synchronous auditory and visual delayed feedback should cause a maximal disruption. On the other hand, if visual feedback is processed independently of auditory feedback, the synchrony should have no effect on the level of disruption. Therefore we included auditory and visual delays that were synchronous (both at 200 ms), and at two levels of asynchrony (DAF: 200 ms and DVF 400 ms and DAF: 200 ms and DVF 600 ms) that exceed the AV integration window.

In addition to investigating the effects of combining delayed auditory and visual feedback on speech fluency, we were also interested in measuring the wide range of changes to the speech signal that DAF has previously been reported to induce. Effects on speech fluency, e.g. increased speech errors and decreased speech rate, are well established. However, the changes in vocal control and rhythm described in early work on DAF have less experimental evidence. Lee (1950) and Fairbanks (1955) noted effects on speech pitch and intensity, and these have been replicated more recently. Under DAF, the intensity of speech (Howell 1990; Stager and Ludlow 1993) and voice pitch (Lechner 1979) increase. A "monotonous" speech pattern has also been described (Fairbanks, 1955; Howell, 2004). However, the only study we are aware of that measured pitch variation under DAF, found no effect in healthy speakers (Brendel et al. 2004).

As well as changing aspects of vocal control, it has been suggested that DAF alters speech rhythm. The 'drawling' (Howell, 2004) and 'monotonous' (Fairbanks 1955; Howell 2004) characteristics of speech under DAF could be accounted for, in part, by disruptions to normal speech rhythm. Howell and Sackin (2002) hypothesized that DAF causes rhythmic disruptions to speech by affecting timing, and demonstrated that the timing of repeated isochronous production of the syllable "ba" was disrupted under DAF (i.e. the variance in syllable timing increased). In the "displaced rhythm hypothesis" Howell and colleagues (Howell et al. 1983; Howell and Archer 1984) argue that speech is just one example of a serially organized behavior that can be disrupted by feedback from an asynchronous rhythmic event (i.e. DAF). In this domain-general explanation of DAF, it is the intensity profile of the delayed speech signal, rather than it's content, that disrupts ongoing speech.

Consistent with this view, DAF of non-speech movements, such as rhythmic finger tapping or keyboard playing (Pfordresher and Benitez 2007), has been shown to impede maintenance of accurate rhythm. In addition, delayed feedback in either the auditory or visual modality slowed production of an isochronous keyboard melody, and this slowing was greatest when both auditory and visual feedback were delayed by the same interval (Kulpa and Pfordresher, 2013). According to Howell and colleague's argument for a domain-independent effect of displaced rhythm on any sequenced action, we could also expect speech production to be disrupted maximally when delaying auditory and visual feedback of speech in synchrony.

Howell and Sackin's (2002) experiment used very simple speech stimuli: a single repeated syllable. To our knowledge, only one previous study has investigated the effect of DAF on rhythm of continuous speech (e.g. words or sentences) (Brendel,et al. 2004), and no significant effect was found. The relative durations of stressed and unstressed syllables contribute to speech rhythm in stress-timed languages such as English. Metrics for rhythm have been developed (Ramus et al. 1999) that compare various measures of vowel or consonant intervals in the speech stream, and these have been applied to cross-linguistic research (White and Mattys 2007) and studies of speech pathologies affecting rhythm (Liss et al. 2009). Two metrics used in these studies that were particularly discriminative were the proportion of an utterance consisting of vowels (%V), and standard deviation of vowel intervals, divided by the mean vowel duration (VarcoV). If vowels are particularly affected under DAF, as has been suggested (Howell, 2004), measures that consider changes in the proportion and variability of vowel intervals would be beneficial for investigating changes in speech rhythm under DAF.

In summary, our study aimed to 1) investigate the effects of combining delayed auditory and visual feedback on the speech of normally fluent people, and 2) provide a comprehensive description of the effects of delayed auditory feedback on speech fluency, vocal control and rhythm.

To address our first aim, we provided speakers with normal or 200 ms delayed auditory feedback, and concurrent visual feedback that was either immediate or delayed (by 200 ms, 400 ms or 600 ms). We hypothesized that the visual feedback would have a stronger effect on speech production when the auditory feedback was delayed than under normal auditory feedback. We used the different levels of visual delay to investigate whether the combination of auditory and visual feedback is maximally disruptive when the delays in the two modalities are synchronous. We expected that immediate visual feedback would decrease the effects of DAF (as suggested by Jones and Striemer, 2007). In contrast, DVF at 200 ms was expected to increase the disruption to speech and to strengthen the effects of DAF at 200 ms. We hypothesized that if DVF influenced speech via integration with DAF, visual delays outside the integration window (400 ms and 600 ms) would not disrupt speech. However, if DVF affects speech independently, DAF combined with DVF at all 3 delay durations would have an effect on speech.

To address our second aim, we compared 200 ms DAF to normal auditory feedback. We predicted that, in line with previous research (Stuart et al. 2002), 200 ms DAF would impair

the fluency of speech. We predicted that sentence duration and speech errors would increase, as found by Stuart and colleagues (2002). In addition, we hypothesized that vocal control would be affected. Specifically, we predicted a change in the intensity of speech, as previously shown by Howell (1990) and Stager and Ludlow (1993), and in pitch, as shown by Lechner (1979). As speech under DAF has previously been described as "monotonous" (Fairbanks, 1955; Howell, 2004), we also predicted that pitch variation would be reduced. Furthermore, we used an automatic speech-to-text alignment method that allowed us to measure changes in speech rhythm (%V and VarcoV). We predicted that the auditory delay would modify speech rhythm. The text-to-speech alignments also allowed us to measure vowel and consonant durations separately and to test the prediction that DAF prolongs vowels specifically. For completeness, we also included the measures of speech rhythm and vocal control in our analyses of the combination of DAF with visual feedback. If a general sensori-motor mechanism for controlling actions is responsible for maintaining rhythmic control of speech, we would expect DVF to disrupt speech rhythm.

II. METHODS

A. Participants

Twenty-two right-handed native English speakers (11 female) took part in the study. The mean age of the participants was 26 years (range: 19-40 years) and they had no history of any communication disorder or neurological impairment. Data of one male participant were excluded from the data analyses, because he was an outlier in majority of the dependent measures (more than 2 standard deviations from the mean for the group). All participants had normal hearing (self-reported) and normal (or corrected-to-normal) vision. The University of Oxford Central University Research Ethics Committee (MSD/IDREC/C1/2011/8) approved this study.

B. Equipment

Participants were presented with audio recordings of sentences using Presentation software (Neurobehavioural Systems) on a Dell desktop computer. They were instructed to repeat each sentence after they had heard it. Participants' speech was recorded using Presentation software and a USB microphone, positioned 40 cm from the participant. An audio signal processor (Alesis Midiverb 4 dual channel signal processor) was used to delay the audio signal of participants' speech by 200 ms in the delayed auditory feedback (DAF) conditions. The audio outputs from the PC and signal processor were passed through a 'Numark M1' 2 channel audio mixer and presented to the participant through Sennheiser HD 280 pro headphones.

In Experiment 1, immediate visual feedback was provided using a 30×30cm mirror, placed at a 60 cm distance from the participant. In Experiment 2, a video camera (Canon Legria HF M32 camcorder) and video signal processor (DelayLine video delay unit, Ovation Systems) were used to delay the video signal of the participant's face. The delayed signal was presented on a 14-in monitor, which was placed 60 cm from the participant, and the camera zoom was adjusted so that the image of the participant's face was the same size as the mirror image of the participant's face in Experiment 1.

C. Stimuli

The stimuli were audio recordings of 20 sentences spoken by a male native speaker of British English. These recording were chosen from a set developed by Davis and colleagues (2011). The sentences had a mean duration of 2137 ms (2002-2306 ms), consisted of 10 syllables (range: 8-11), and were matched for number of labially produced phonemes (which are more easily visually discriminated).

D. Procedure

Participants completed two separate experiments: experiment 1 investigated the effects of immediate visual feedback and delayed auditory feedback on speech production, and experiment 2 examined the effects of delayed visual feedback and delayed auditory feedback on speech production. The procedure was organized in this way for practical reasons, as immediate feedback was given using a mirror, whilst delayed visual feedback was displayed electronically, as described in section B above. Conditions that acted as baselines were repeated in experiments 1 and 2. All participants completed experiment 1 before experiment 2.

In Experiment 1 participants repeated all 20 sentences in each of 4 experimental conditions:

1) Immediate auditory feedback and no visual feedback (IAF+NVF)

2) Immediate auditory feedback and immediate visual feedback (IAF+IVF)

3) Delayed auditory feedback and no visual feedback (DAF+NVF)

4) Delayed auditory feedback and immediate visual feedback (DAF+IVF)

The order of the conditions was pseudo-randomized: 4 blocks of 5 consecutive sentences for each condition were presented, and the order of these blocks was randomized. Participants heard each sentence, and then repeated it over pink noise (noise with equal energy per octave). The purpose of the noise was to mask any immediate auditory feedback in DAF conditions in which the auditory feedback was delayed by 200 ms. In the IVF conditions, participants saw themselves speaking in the mirror, and in the NVF conditions the mirror was covered and participants were instructed to attend to a fixation cross at an equivalent position to their mouth when viewing their mirror image. Participants were not explicitly instructed to attend to the auditory feedback, or to control their speech production in any way.

In Experiment 2, the participants repeated the same 20 sentences in each of eight experimental conditions:

1) Immediate auditory feedback and no visual feedback (IAF+NVF)

2) Immediate auditory feedback and 200ms delayed visual feedback (IAF +DVF200)

3) Immediate auditory feedback and 400ms delayed visual feedback (IAF +DVF400)

4)   Immediate auditory feedback and 600ms delayed visual feedback (IAF +DVF600)

5)   Delayed auditory feedback and no visual feedback (DAF+NVF)

6)   Delayed auditory feedback and 200ms delayed visual feedback (DAF+DVF200)

7)   Delayed auditory feedback and 400ms delayed visual feedback (DAF+DVF400)

8)   Delayed auditory feedback and 600ms delayed visual feedback (DAF+DVF600)

As in Experiment 1, the order of the conditions was pseudo-randomized, and pink noise was presented during production of the sentences. In all conditions including DAF, the auditory feedback was delayed by 200 ms. In the DVF conditions participants viewed their speaking faces on the monitor, and were instructed to attend to their mouth. In the NVF conditions the monitor was covered, and participants were instructed to attend to a fixation cross aligned to the position of the image of their mouth in the DVF conditions. Again, participants were not explicitly instructed to attend to the auditory feedback, or to control their speech production in any way.

E. Speech measures

Measures of duration, intensity, mean and standard deviation of pitch of participants' speech recordings were determined using automated scripts in Praat (Boersma and Weenink 1992-2002). Durations were checked and adjusted manually. Speech errors were counted manually. Speech errors were defined as repetitions, omissions or substitutions (on a phoneme, syllable or word-level). The first rater counted errors for all participants, and a second rater counted errors for a subset of 15 participants. Both raters were blind to the condition during rating. Inter-rater agreement was calculated using Pearson's product moment correlation co-efficient, and the raters' scores were significantly correlated ($r = .97$, $n = 180$, $p < .0005$). The speech error counts from the first rater were used for the analysis.

In order to measure changes in speech rhythm, resulting from the feedback manipulations, we transcribed participant's speech and performed a text-to-speech alignment. The alignment system included acoustic-phonetic models from the "Penn Phonetics Forced Aligner", p2fa (Yuan and Liberman 2008), with a customized lexicon with appropriate British English phonemes substituted for their North American equivalents. This lexicon included alternative pronunciations of words where variations due to dialect or speaking style are common. Viterbi alignment was performed using the Hidden Markov Model Toolkit, HTK.

The phoneme models were standard HTK-format 3-state, left-to-right, monophone models. The front-end produced 12th-order Perceptual Linear Prediction coefficients, including the zeroth, together with delta and delta-delta coefficients. The frame-rate was one every 10 ms, with a 25 ms Hamming analysis window. There were separate models for initial/final silence, short inter-word pauses, and various non-speech sounds.

The alignments were manually checked, and data from 5 participants were rejected due to insufficient accuracy. These mis-alignments were most likely caused by strong regional accents. The text-to-speech alignment enabled us to calculate two measures of rhythm: the

percentage of the each sentence duration consisting of vowels (%V: see Ramus et al. (1999) for full description), and the standard deviation of vowel durations divided by mean vowel duration (VarcoV: see White and Mattys, 2007 for full description). In addition, the alignments allowed us to obtain durations of vowel and consonant segments.

F. Statistical analyses

To assess the effects of DAF on speech production we calculated DAF 'change scores'. For each speech measure, the scores for the DAF+NVF conditions were subtracted from the scores for IAF+NVF conditions. The scores were averaged across experiments 1 and 2 (since no differences were found between experiments in the scores for the DAF+NVF or IAF+NVF conditions). The effects of DAF on each speech measure were statistically tested using one sample t-tests (two-tailed).

Data from Experiment 1 were used to assess the effects of immediate visual feedback on speech production. We calculated IVF 'change scores' for each speech measure, by subtracting the scores for NVF condition from the scores for IVF condition at both levels of auditory feedback (i.e. (IAF+IVF)-(IAF+NVF) and (DAF+IVF)-(DAF+NVF)). We used one-sample t-tests (two-tailed) to test the effect of IVF on each measure under DAF and IAF.

Data from Experiment 2 were used to assess the effects of DVF on speech production. First, we calculated DVF 'change scores' for each speech measure by subtracting the scores for NVF conditions from the scores for DVF conditions at both levels of auditory feedback (i.e. (IAF+DVF200)-(IAF+NVF), (DAF+DVF200)-(DAF+NVF), etc.). We then performed a 2×3 repeated measures ANOVA for each speech measure separately, with auditory feedback condition (immediate, delayed) and the level of the visual delay (200 ms, 400 ms, 600 ms) as within-subject factors. These analyses tested whether the different levels of visual delay affected speech, and how these effects interacted with DAF. When a significant interaction was found we carried out one-way ANOVAs for the auditory feedback levels separately. Two-tailed paired t-tests were then used for comparisons between levels of visual delay. In addition, we tested whether the visual delay conditions differed from the no visual feedback baseline conditions, using two-tailed one-sample t-tests. When there was no significant difference between the levels of visual delay we performed the t-test on the average of the 3 levels of visual delay.

III. RESULTS

A. Effects delayed auditory feedback on speech

The effects of DAF on each of the speech measures are presented in Table I. All measures were significantly affected by DAF. The measures of speech fluency were all significantly changed: Sentence durations were significantly longer, and there were more speech errors under DAF than IAF. Durations of vowel and consonant segments were also both prolonged; there was no significant difference between them. In addition, vocal control was affected by DAF: Mean pitch and pitch variation were reduced, and mean speech intensity

increased under DAF. Also, DAF affected both measures of speech rhythm: V% and VarcoV.

B. Effects of immediate visual feedback on speech

The effects of IVF on each of the speech measures are presented in Table II. IVF prolonged consonant durations when combined with IAF, but not when combined with DAF. IVF had no effect on any of the other measures of speech fluency, when combined with IAF or DAF. Furthermore, IVF had no effect on speech intensity, mean pitch, pitch variation or either measure of speech rhythm (%V, VarcoV).

C. Effects of delayed visual feedback on speech

For each speech measure, the average effect of DVF, compared to the no-visual feedback baseline, is presented in Table III. The detailed results for the measures that were significantly affected by the level of visual delay are presented in Fig. 1. DVF increased sentence durations when combined with DAF (a significant main effect of auditory feedback condition, $F_{1,20} = 15.52$, $p = .001$; Fig. 1a; and a significant increase from NVF, Table III). DVF had no effect on sentence durations when combined with IAF (Table III). The duration of visual delay had no effect on sentence durations (no significant main effect of visual delay level, no significant interaction between visual delay level and auditory feedback condition).

DVF increased the number of speech errors under DAF (a significant main effect of auditory feedback condition: $F_{1,20} = 6.37$, $p = .02$; Fig. 1b; and a significant increase from NVF, Table III). The duration of the visual delay also influenced the number of errors (a significant main effect of visual delay level: $F_{2,40} = 3.33$, $p = .046$; a significant interaction between auditory feedback condition and visual delay level: $F_{2,40} = 8.94$, $p = .001$). The duration of visual delay significantly affected the number of errors under DAF ($F_{2,40} = 6.79$, $p = .003$; Fig. 1b), but had no effect under IAF. Under DAF, the number of speech errors was increased when visual feedback was delayed by 200 ms ($t_{20} = 2.92$, $p = .008$) and by 400 ms ($t_{20} = 2.94$, $p = .008$) compared to when no visual feedback was given (see Fig. 1b). However, the visual delay of 600 ms had no effect on the number of speech errors. There was also a significant difference between the number of errors made when visual feedback was delayed by 200 ms compared to 600 ms ($t_{20} = 3.30$, $p = .004$), and with 400 ms compared to 600 ms ($t_{20} = 3.66$, $p = .002$).

Vowel durations were affected by DVF when combined with DAF (a significant main effect of auditory feedback condition: $F_{1,15} = 8.83$, $p = .01$; Fig. 1c, and a significant increase from NVF Table III). DVF had no effect on vowel durations when combined with IAF (Table III). The duration of visual delay had no effect on vowel durations (no significant main effect of visual delay level, no significant interaction between visual delay level and auditory feedback condition). Consonant durations were not significantly affected by DVF (no significant main effects or interactions; Table III).

Most aspects of vocal control (speech intensity and pitch variation) were not significantly affected by DVF when combined with IAF or DAF (no significant main effects or interactions; Table III). The mean pitch was decreased by DVF when combined with DAF,

but not when combined with IAF (Table III). This difference between IAF and DAF conditions was not, however, significant (no significant main effects or interactions involving auditory feedback condition).

The %V measure increased when DVF was combined with DAF (main effect of auditory feedback level: $F_{1,15} = 6.29$, $p = .024$; Fig. 1d; and a significant increase from NVF, Table III). DVF had no effect on %V when combined with IAF (Table III). The duration of the delay had no effect on %V (no significant main effect of visual delay level, no significant interaction between visual delay level and auditory feedback condition). The vowel variation measure (VarcoV) was not significantly affected by DVF, when combined with IAF or DAF (no significant main effects or interactions, Table III).

IV. DISCUSSION

The main aim of the current study was to investigate whether DVF affects speech production, since this has not previously been investigated in fluent speakers. We examined effects of both DVF and IVF on speech production in the presence of IAF and DAF. We also manipulated the synchrony between visual and auditory feedback, to test whether visual feedback influences speech production via integration with auditory feedback. In addition, we considered the range of effects on speech previously ascribed to DAF. Although the effects of DAF on fluency have been well documented, the effects on vocal control and rhythm have very little experimental evidence. In the current study, we performed a detailed speech analysis that allowed us to determine effects of DVF, IVF and DAF on speech fluency (i.e., speech rate and errors), vocal control (i.e., intensity and pitch) and speech rhythm (%V, VarcoV).

We first discuss the effects of DAF. Our study replicated previous findings regarding disruption of fluency (e.g. Fairbanks and Guttman 1958; Yates 1963). We found that DAF increased both sentence duration and the number of speech errors made. In addition, we separately measured the changes in the durations of consonant and vowel segments of speech under DAF. Early descriptions of speech under DAF likened it to stuttering as both may induce phoneme repetitions and prolongations, or pauses in the speech stream. However, this parallel has some limitations. Several authors have described that stuttering particularly affects consonants (Jayaram 1983; Howell et al. 1988), and mostly those occurring at the beginnings of words and sentences (Wingate, 2002). By contrast, Howell (2004) argued that DAF affects 'medial vowels' (which we take to mean vowels within the syllable nucleus) most strongly. We found that both vowel and consonant segments were prolonged under DAF. Therefore, our results suggest that speech under DAF differs in this respect from stuttered speech. However, it should be noted that the distinctions previously made between DAF speech and stuttering refer also to phoneme position, and to all types of speech errors (repetitions and pauses, in addition to prolongations). In the current study, we did not examine the effect of phoneme position on either prolongations or speech errors.

In agreement with previous studies (Howell 1990; Stager and Ludlow 1993) our findings showed that DAF increased intensity of speech. We also found a decrease in pitch variation, which supports the observations that speech sounds more "monotonous" under DAF

(Fairbanks, 1955; Howell, 2004). However, in contrast to the early studies (Fairbanks, 1955; Lechner, 1979) that reported that mean pitch increases under DAF, we found that mean pitch was decreased under DAF. Fairbanks (1955) suggests that the increase in pitch may be caused by increased muscular tension associated with the participants' attempts to "resist experimental interference". It might be the case that participants actively, rather than passively, modify voice pitch in order to overcome the effects of DAF. The participants in Fairbanks' (1955) and Lechner's (1979) studies were all male, and Fairbanks reports a lower mean pitch in the IAF condition than in our study, which included equal numbers of male and female participants. One possibility is that our participants used the strategy of lowering rather than increasing pitch under DAF, due to a relatively higher natural speech pitch.

It has previously been suggested that DAF disrupts the rhythm of speech. Howell and Sackin (2002) demonstrated this during production of regular single syllables. We found that both the VarcoV and %V rhythm measures were increased under DAF, when our participants repeated full sentences. One previous study calculated rhythmic speech changes during continuous speech, but no changes in rhythm were found (Brendel et al. 2004). That study used a Pairwise Variability Index (PVI) for vowels to measure rhythm: the mean of the difference between successive vowel intervals divided by their sum. The authors noted that further work would be needed to find out whether this is the most appropriate metric to measure change in rhythm under altered feedback. Although PVI is a reliable metric, Liss and White (2009) found %V and VarcoV to be better able to distinguish between dysarthria subtypes, and between dysarthric and normal speakers than the PVI measure used by Brendel and colleagues. Such sensitivity to pathological aspects of speech seems most relevant to our consideration of rhythm change under DAF.

Regarding the visual feedback effects, Jones and Striemer (2007) found that IVF can reduce the disruptive effects of DAF. We failed to replicate these findings; we found no reduction in speech errors when IVF was combined with DAF, compared to when no visual feedback was given. IVF had no effects on measures of speech rhythm or vocal control either. This lack of replication is not surprising as Jones and Striemer (2007) only found effects when they divided their experimental group according to response to DAF alone; only the 'low disruption' group benefitted from the visual feedback. However, we did find a prolongation of consonant segments when IVF was combined with IAF (relative to IAF alone). It is difficult to interpret this finding, as the effect was absent when IVF was combined with DAF. Further work is needed to find out whether this effect is replicable.

The main aim of the current study was to investigate whether DVF affects speech production in normally fluent speakers. We hypothesized that delayed visual feedback may influence speech production, either through integration with auditory information within the auditory feedback mechanism described in the DIVA model, or via an independent feedback route. We found that DVF had no effects on speech production when auditory feedback was not delayed. However, when auditory feedback was delayed, DVF increased the effects of DAF on speech fluency (i.e. durations and errors), in line with our prediction. In addition, we found an effect on the %V measure of speech rhythm, which increased when DVF and DAF were combined. However, our other measure of rhythm, VarcoV, was not affected by DVF.

There is some previous evidence for a hierarchy of the effects of sensory feedback on speech. For example, auditory feedback has greater effects on the fluency of stutterers, than visual feedback (Kalinowski and Stuart 1996; Kalinowski et al. 2000). Auditory feedback also dominates visual feedback in the timing of rhythmic manual actions (Pfordresher and Palmer 2002; Repp and Penel 2002; Repp and Penel 2004). This would indicate that auditory feedback influences the production of motor actions more strongly than visual feedback. It should be also noted that visual feedback from speech articulatory movements is not usually available to a speaker, so this information is unlikely to be heavily weighted. Hence, if auditory feedback indicates that no errors are being made in speech, delayed visual feedback alone may not be strong enough to generate an error signal. However, when the auditory feedback is perturbed, an error signal would be generated, and visual feedback may make an additional contribution to an already de-stabilized speech system. Our finding is inconsistent with the study by Kulpa and Pfordesher (2013), which found that DVF alone could disrupt a manual keyboard melody task. In contrast to speech, visual feedback from manual actions is usually available, and it may, therefore, have a greater contribution to motor control.

The duration of the visual delay had no major effect in the current study. We used a fixed auditory delay of 200 ms, which previous studies in fluent speakers have shown to be maximally disruptive to speech (Stuart et al. 2002) and visual delays of 200, 400 and 600 ms. We hypothesized that a greater disruption to speech production when auditory and visual delayed feedback were in synchrony would indicate that feedback from the two modalities integrate during speech production. However, we found no significant difference in most speech measures between the 3 delay levels. One exception was number of speech errors, which did further increase when DAF was combined with DVF of 200 or 400 ms, but not 600 ms. The finding that visual delays of up to 600 ms could induce speech disruptions when paired with 200 ms DAF is inconsistent with the idea that audio-visual feedback would create an integrated error signal, since integration is unlikely to occur at such a level of asynchrony. Therefore, we suggest that visual feedback is processed separately from auditory feedback during speech production.

Our results indicated that the fluency measures, but not voice control and rhythm measures, were sensitive to the combination of delayed auditory and visual feedback. Speech information contained in the visual signal relates to articulatory movements during speech production (Summerfield, 1992), so it would follow that the articulation of speech (i.e., fluency) could be affected by a visual feedback delay. However, intensity and pitch cues are not likely to be visible on a speaker's face and, therefore, it makes sense that these aspects of speech production are not disrupted by DVF. We did find, however, that one of the two rhythm measures (%V) was also sensitive to the combination of delayed AF and VF. %V is not normalized for speech rate, unlike our other rhythm measure (VarcoV), which showed no significant effects of DVF. Dellwo and Wagner (2003) have argued that %V is quite robust to rate changes, as consonant and vowel duration tend to change in comparable degrees when rate is altered. However, it may be the case that DVF did increase the proportion of the speech sample consisting of vowels, but did so equally for all vowels, resulting in no overall change to speech rhythm (and so no effect on VarcoV). Since DVF

increased speech rate and had no effect on the rate-normalized measure of rhythm (VarcoV), we conclude that DVF did not modulate speech rhythm.

Our finding of a lack of effect of DVF on rhythm is inconsistent with experiments showing an effect of the combination of DAF and DVF on non-speech actions (Kulpa and Pfordresher 2013), which indicates that the rhythmic control of speech is maintained by a speech-specific system. It should be noted that the combination of DAF and DVF in Kulpa and Pfordresher's (2013) experiment disrupted performance of the keyboard melody by a combination of increased task duration and timing variability. We would consider only the timing aspect of this effect a disruption to rhythm. In their experiment, DAF contributed to duration but not timing variability (Pfordresher and Palmer 2002), and DVF increased both measures. We also found that the combination of DAF and DVF slowed speech, but the timing variability was increased by DAF alone (VarcoV). These differences in the relative contributions of DAF and DVF likely relate to how sensory feedback from each modality is used in speech and non-speech domains. This would be consistent with separate system for speech and non-speech rhythmic control. However, a possible alternative would be a unitary mechanism that uses different weightings for each sensory modality, in speech and non-speech domains.

In sum, we found that DAF affected fluency, vocal control and rhythm of speech, indicating that perturbations of auditory feedback can disrupt all aspects of speech production. DVF strengthened the DAF-induced disruptions to fluency but not to rhythm and vocal control. This suggests that although speakers do not normally see their articulatory movements, they can use visual feedback about these movements when it is available. The finding that DVF does not need to be synchronous with DAF in order to disrupt fluency suggests that visual feedback influences speech production independently, rather than through integration with the auditory signal.

ACKNOWLEDGEMENTS

References

Boersma P, Weenink D. Praat, a system for doing phonetics by computer v.4.0.13. 1992-2002

Brendel B, Lowit A, Howell P. The effects of delayed and frequency shifted feedback on speakers with Parkinson's Disease. J Med Speech Lang Pathol. 2004; 12:131–138. [PubMed: 19330038]

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS. Activation of auditory cortex during silent lipreading. Science. 1997; 276(5312):593–596. [PubMed: 9110978]

Civier O, Tasko SM, Guenther FH. Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production. J Fluency Disord. 2010; 35(3):246–279. [PubMed: 20831971]

Davis MH, Ford MA, Kherif F, Johnsrude IS. Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI. Journal of Cognitive Neuroscience. 2011; 23(12):3914–3932. [PubMed: 21745006]

Dellwo V, Wagner P. Relations between language rhythm and speech rate. Proceedings of the 15th international congress of phonetics sciences. 2003:471–474.

Elman JL. Effects of frequency-shifted feedback on the pitch of vocal productions. The Journal of the Acoustical Society of America. 1981; 70(1):45–50. [PubMed: 7264071]

Fairbanks G. Selective Vocal Effects Of Delayed Auditory Feedback. Journal of Speech and Hearing Disorders. 1955; 20(4):333–346. [PubMed: 13272227]

Fairbanks G, Guttman N. Effects of delayed auditory feedback upon articulation. Journal of Speech and Hearing Research. 1958; 1(1):12–22. [PubMed: 13576494]

Fridriksson J, Hubbard HI, Hudspeth SG, Holland AL, Bonilha L, Fromm D, Rorden C. Speech entrainment enables patients with Broca's aphasia to produce fluent speech. Brain: a journal of neurology. 2012; 135(Pt 12):3815–3829. [PubMed: 23250889]

Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 2006; 96(3):280–301. [PubMed: 16040108]

Houde JF, Jordan MI. Sensorimotor adaptation in speech production [see comments]. Science. 1998; 279(5354):1213–1216. [PubMed: 9469813]

Howell P. Changes in Voice Level Caused by Several Forms of Altered Feedback in Fluent Speakers and Stutterers. Language and Speech. 1990; 33:325–338. [PubMed: 2133911]

Howell P. Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering. Stammering Res. 2004; 1(1):31–46. [PubMed: 18259594]

Howell P, Archer A. Susceptibility to the effects of delayed auditory feedback. Perception & Psychophysics. 1984; 36(3):296–302. [PubMed: 6522222]

Howell P, Powell DJ, Khan I. Amplitude Contour of the Delayed Signal and Interference in Delayed Auditory-Feedback Tasks. Journal of Experimental Psychology-Human Perception and Performance. 1983; 9(5):772–784.

Howell P, Sackin S. Timing interference to speech in altered listening conditions. J Acoust Soc Am. 2002; 111(6):2842–2852. [PubMed: 12083218]

Howell, P.; Wingfield, T.; Johnson, M.; Ainsworth, WAH.; N., J.; Lawrence, R. Speech '88: Proceedings. Institute of Acoustics; Edinburgh: 1988. Characteristics of the Speech of Stutterers During Normal and Altered Auditory Feedback; p. 1069-1076.

Hudock D, Dayalu VN, Saltuklaroglu T, Stuart A, Zhang J, Kalinowski J. Stuttering inhibition via visual feedback at normal and fast speech rates. International journal of language & communication disorders / Royal College of Speech & Language Therapists. 2010

Jayaram M. Phonetic influences on stuttering in monolingual and bilingual stutterers. Journal of communication disorders. 1983; 16(4):287–297. [PubMed: 6571179]

Jones JA, Munhall KG. Perceptual calibration of F0 production: evidence from feedback perturbation. The Journal of the Acoustical Society of America. 2000; 108(3 Pt 1):1246–1251. [PubMed: 11008824]

Jones JA, Munhall KG. Remapping auditory-motor representations in voice production. Current biology. 2005; CB 15(19):1768–1772. [PubMed: 16213825]

Jones JA, Striemer D. Speech disruption during delayed audiotry feedback with simultaneous visual feedback. Journal of the Acoustical Society of America. 2007; 122(4):135–141. [PubMed: 17614473]

Kalinowski J, Stuart A. Stuttering amelioration at various auditory feedback delays and speech rates. European journal of disorders of communication: the journal of the College of Speech and Language Therapists, London. 1996; 31(3):259–269.

Kalinowski J, Stuart A, Rastatter MP, Snyder G, Dayalu V. Inducement of fluent speech in persons who stutter via visual choral speech. Neuroscience Letters. 2000; 281(2-3):198–200. [PubMed: 10704777]

Kuhl PK, Meltzoff AN. The bimodal perception of speech in infancy. Science. 1982; 218(4577):1138–1141. [PubMed: 7146899]

Kulpa JD, Pfordresher PQ. Effects of delayed auditory and visual feedback on sequence production. Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale. 2013; 224(1):69–77.

Lechner BK. Effects of Delayed Auditory-Feedback and Masking on the Fundamental-Frequency of Stutterers and Non-Stutterers. Journal of Speech and Hearing Research. 1979; 22(2):343–353. [PubMed: 491561]

Lee BS. Effects of Delayed Speech Feedback. Journal of the Acoustical Society of America. 1950; 22(6):824–826.

Lee BS. Some Effects of Side-Tone Delay. Journal of the Acoustical Society of America. 1950; 22(5): 639–640.

Liss JM, White L, Mattys SL, Lansford K, Lotto AJ, Spitzer SM, Caviness JN. Quantifying speech rhythm abnormalities in the dysarthrias. Journal of speech, language, and hearing research: JSLHR. 2009; 52(5):1334–1352.

McGurk H, MacDonald J. Hearing lips and seeing voices. Nature. 1976; 264:746–748. [PubMed: 1012311]

Mottonen R, Krause CM, Tiippana K, Sams M. Processing of changes in visual speech in the human auditory cortex. Brain Research. Cognitive Brain Research. 2002; 13(3):417–425. [PubMed: 11919005]

Pfordresher PQ, Benitez B. Temporal coordination between actions and sound during sequence production. Human movement science. 2007; 26(5):742–756. [PubMed: 17764772]

Pfordresher PQ, Palmer C. Effects of delayed auditory feedback on timing of music performance. Psychological research. 2002; 66(1):71–79. [PubMed: 11963280]

Ramus F, Nespor M, Mehler J. Correlates of linguistic rhythm in the speech signal. Cognition. 1999; 73(3):265–292. [PubMed: 10585517]

Repp BH, Penel A. Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. Journal of experimental psychology. Human perception and performance. 2002; 28(5):1085–1099. [PubMed: 12421057]

Repp BH, Penel A. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. Psychological research. 2004; 68(4):252–270. [PubMed: 12955504]

Rosenbek JC, Lemme ML, Ahern MB, Harris EH, Wertz RT. A treatment for apraxia of speech in adults. The Journal of speech and hearing disorders. 1973; 38(4):462–472. [PubMed: 4754879]

Shiller DM, Sato M, Gracco VL, Baum SR. Perceptual recalibration of speech sounds following speech motor learning. The Journal of the Acoustical Society of America. 2009; 125(2):1103–1113. [PubMed: 19206885]

Siegel GM, Pick HL Jr. Auditory feedback in the regulation of voice. The Journal of the Acoustical Society of America. 1974; 56(5):1618–1624. [PubMed: 4427032]

Snyder GJ, Hough MS, Blanchet P, Ivy LJ, Waddell D. The effects of self-generated synchronous and asynchronous visual speech feedback on overt stuttering frequency. Journal of Communication Discorders. 2009; 42:235–244.

Stager SV, Ludlow CL. Speech Production Changes under Fluency-Evoking Conditions in Nonstuttering Speakers. Journal of Speech and Hearing Research. 1993; 36(2):245–253. [PubMed: 8487517]

Stuart A, Kalinowski J, Rastatter MP, Lynch K. Effect of delayed auditory feedback on normal speakers at two speech rates. The Journal of the Acoustical Society of America. 2002; 111(5 Pt 1): 2237–2241. [PubMed: 12051443]

Sumby WH, Pollack I. Visual contribution to speech intelligibility in noise. The Journal of the Acoustical Society of America. 1954; 26(2):212–215.

Summerfield Q. Lipreading and audio-visual speech perception. Phil Trans R Soc Lond B Biol Sci. 1992; 335(1273):71–78. [PubMed: 1348140]

Tourville JA, Guenther FH. The DIVA model: A neural theory of speech acquisition and production. Language and Cognitive Processes. 2011; 26(7):952–981. [PubMed: 23667281]

Tye-Murray N. Visual feedback during speech production. Journal of the Acoustical Society of America. 1986; 79(4):1169–1171. [PubMed: 3700869]

van Wassenhove V, Grant KW, Poeppel D. Temporal window of integration in auditory-visual speech perception. Neuropsychologia. 2007; 45(3):598–607. [PubMed: 16530232]

Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastian-Galles N, Werker JF. Visual language discrimination in infancy. Science. 2007; 316(5828):1159. [PubMed: 17525331]

White L, Mattys SL. Calibrating rhythm: First language and second language studies. Journal of Phonetics. 2007; 35(4):501–522.

Williams, P.; Bowen, C. Children's speech sound disorders. Wiley-Blackwell; Oxford: 2009. The Nuffield approach to childhood apraxia of speech and other severe speech disorders; p. 270-275.

Wingate, ME. Foundations of stuttering. Academic Press; San Diego: 2002.

Yates AJ. Delayed auditory feedback. Psychological bulletin. 1963; 60:213–232. [PubMed: 14002534]

Yuan J, Liberman M. Speaker identification on the SCOTUS corpus. Proceedings of Acoustics. 2008; 2008:5687–5690.
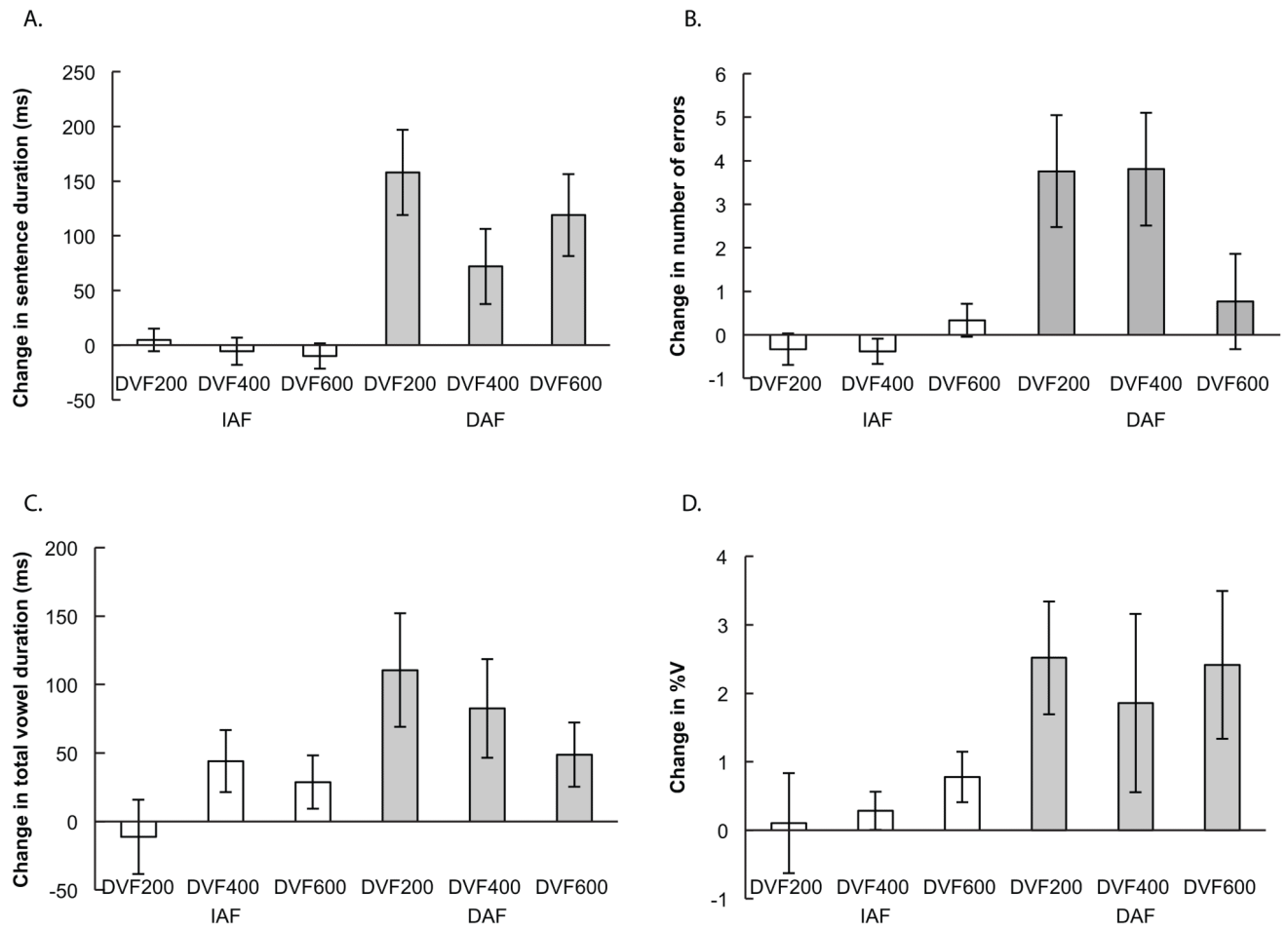
Figure 1.
Effects of delayed visual feedback on speech. The graphs present mean changes (± standard error) relative to the no visual feedback condition, for each of the delayed visual feedback conditions (DVF200: 200 ms delayed visual feedback, DVF400: 400 ms delayed visual feedback, DVF600: 600 ms delayed visual feedback) combined with immediate auditory feedback (IAF) or delayed auditory feedback (DAF). A) Sentence durations, B) Speech errors, C) Total vowel segment duration per sentence, D) %V measure.

TABLE I

The effect of delayed auditory feedback on speech production.

Speech measure	n	IAF (s.e.m.)	DAF (s.e.m.)	Mean change (s.e.m.)	t (p value)
Fluency					
Sentence duration (ms)	21	2171 (46)	3128 (218)	957.17 (200.26)	**4.78 (.001)**
Total no. of speech errors	21	2.29 (0.57)	15.29 (2.03)	13 (1.85)	**7.04 (.001)**
Total duration of consonant segments (ms)	16	1087 (33)	1271 (77)	183.93 (67.86)	**2.71 (.016)**
Total duration of vowel segments (ms)	16	841 (42)	1106 (113)	265.14 (101.57)	**2.61 (.020)**
Vocal control					
Mean intensity (dBSPL)	21	58.63 (077)	61.82 (0.68)	3.19 (0.3)	**10.47 (.001)**
Mean pitch (Hz)	21	222.35 (7.52)	206.88 (7.94)	−15.46 (4.89)	**−3.16 (.005)**
Standard deviation of pitch (Hz)	21	111.00 (9.90)	91.72 (8.49)	−19.28 (4.52)	**−4.26 (.001)**
Rhythm					
VarcoV	16	62.12 (1.26)	66.73 (1.77)	4.62 (1.42)	**3.25 (.005)**
%V	16	43.28 (0.79)	45.74 (1.22)	2.45 (0.87)	**2.82 (.013)**

TABLE II

The effect of immediate visual feedback on speech production.

Speech measure	IAF + IVF (s.e.m)	Mean change* (s.e.m)	t (p value)	DAF + IVF (s.e.m)	Mean change* (s.e.m)	t (p value)
Fluency						
Sentence duration (ms)	2206 (46)	27.07 (14.34)	1.89 (ns.)	3071 (196)	−32.55 (57.58)	−.57 (ns)
Total no. of speech errors	3.62 (0.99)	1.09 (.71)	1.54 (ns)	14.86 (1.74)	−0.05 (1.18)	−.04 (ns)
Total duration of consonant segments (ms)	1124 (34)	37.14 (15.91)	**2.33 (.034)**	1282 (84)	10.47 (42.77)	.25 (ns.)
Total duration of vowel segments (ms)	855 (39)	13.90 (11.49)	1.21 (ns.)	1149 (115)	42.63 (21.40)	1.99 (ns.)
Vocal control						
Mean intensity (dBSPL)	58.43 (0.73)	−0.29 (0.27)	−1.06 (ns)	61.70 (0.81)	−0.30 (0.18)	−1.62 (ns)
Mean pitch (Hz)	219.46 (7.95)	−1.93 (2.48)	−.78 (ns)	204.16 (7.60)	−1.83 (2.00)	−.91 (ns)
Standard deviation of pitch (Hz)	110.43 (9.34)	−1.59 (2.21)	−.72 (ns.)	90.45 (8.62)	−1.93 (2.43)	−.79 (ns.)
Rhythm						
VarcoV	61.76 (1.96)	−0.22 (1.04)	−.22 (ns)	65.90 (2.15)	−1.89 (1.32)	−1.43 (ns)
%V	42.96 (0.74)	−0.09 (.39)	−.22 (ns)	46.69 (1.41)	0.32 (0.60)	.54 (ns)

*
Mean change relative to no visual feedback condition.

TABLE III

The effect of delayed visual feedback on speech production. The average value for the 3 levels of visual delay is reported.

Speech measure	n	IAF + DVF (s.e.m)	Mean change * (s.e.m)	t (p value)	DAF + DVF (s.e.m)	Mean change * (s.e.m)	t (p value)
Fluency							
Sentence duration (ms)	21	2160 (51)	−3.61 (11.44)	−0.35 (ns)	3269 (259)	116.31 (36.96)	**4.06 (.001)**
Total no. of speech errors	21	1.92 (0.42)	−0.12 (0.34)	−0.50 (ns)	18.45 (2.90)	2.78 (1.23)	**2.52 (.020)**
Total duration of consonant segments (ms)	16	1084 (38)	−3.13 (16.58)	−0.33 (ns.)	1271 (85)	−0.07 (35.68)	−0.00 (ns.)
Total duration of vowel segments (ms)	16	862 (48)	20.5 (23.05)	1.27 (ns)	1187 (133)	80.60 (33.63)	**2.68 (.017)**
Vocal control							
Mean intensity (dBSPL)	21	58.83 (0.77)	0.28 (0.23)	1.33 (ns)	61.84 (0.70)	0.18 (0.19)	1.16 (ns)
Mean pitch (Hz)	21	225.20 (8.71)	1.9 (2.74)	0.81 (ns)	203.66 (8.10)	−4.13 (1.47)	**−2.85 (.010)**
Standard deviation of pitch (Hz)	21	111.43 (10.10)	1.44 (2.26)	0.78 (ns)	87.61 (8.82)	−3.44 (2.60)	−1.53 (ns)
Rhythm							
VarcoV	16	62.22 (1.41)	−0.03 (1.17)	−0.03 (ns)	67.19 (2.05)	1.52 (1.22)	1.49 (ns)
%V	16	43.91 (0.69)	0.39 (0.46)	1.09 (ns.)	47.38 (0.96)	2.26 (107)	**2.27 (.038)**

*
Mean change relative to no visual feedback condition.