# Impact of Altered Sensory Feedback on Speech Control in Fluent Speakers and Speakers who Stutter

Liam Barrett and Peter Howell*

Experimental Psychology, University College London

*Send correspondence to: Peter Howell, Experimental Psychology, University College London, Gower Street, London WC1E 6BT (e-mail: p.howell@ucl.ac.uk).

# Introduction

During speech control, the brain generates a motor plan for an utterance and issues it to the articulators. Sensory consequences occur as the motor plan is performed that provide information about the utterance. This information can be returned over afferent pathways to allow discrepancies between the planned (efferent) and output (afferent) forms of speech to be detected. Monitoring for discrepancies would allow planning errors to be identified and feedback control occurs when errors are corrected (Howell, 2004a). The afferent information is carried in several sensory modalities (e.g. auditory, kinaesthetic, and somatosensory). Audition is the modality used most often in feedback theories where the idea is that speakers listen to their speech output whilst it is being produced and speech is interrupted and corrected when planning errors are heard. A way of testing such theories is to perturb the auditory signal (change the temporal, spectral or intensity properties) that makes it appear that an error has occurred during articulation. Speakers then respond to correct the putative error and this disrupts speech control. Hence, speech disruption under auditory perturbations is considered to support auditory feedback theories.

Superficially, the disruptions to speech under auditory perturbations are similar to the dysfluencies that people who stutter (PWS) show in their unperturbed speech. This could arise because the correction-process that works for fluent speakers malfunctions in PWS. Furthermore, when the auditory feedback (AFB) of PWS was perturbed, fluency was enhanced (Howell, 2004a for review). This was interpreted as showing that the problem PWS had in dealing with auditory feedback was corrected when the afferent signal was manipulated. The research on perturbations led to neural accounts of fluency control of speech, whilst the studies with PWS was exploited in interventions for speech control problems (Costello-Ingham, 1993, p.30; Goldiamond 1965; Howell, 2004b; Ryan, 1974; Tourville & Guenther, 2011). This chapter reviews the effects temporal, spectral and intensity perturbations to auditory and vibratory feedback have on fluent speakers and PWS.

# Altered Sensory Feedback

Procedures to alter the sensory consequences of own speech output include changes to AFB and VibroTactile Feedback (VTF). With AFB and VTF, temporal, spectral and intensity parameters can be altered separately or together. Furthermore, the sensory parameters have been manipulated in different ways (e.g. the spectrum of all, or just some, components of speech have been altered) and their impact on speech control may differ.

# Temporal changes to AFB (Delayed Auditory Feedback)

## Fluent speakers

Speech can be delayed relative to normal listening conditions, referred to as Delayed Auditory Feedback (DAF). Early work on DAF with fluent speakers, showed that speech errors[1] arise (Black, 1951; Fairbanks, 1955; Lee, 1950), speakers increase voice level (Howell, 1990), speech is slowed, medial vowels in syllables are elongated (Howell, Wingfield & Johnson, 1988) and pitch is monotone. The way speakers respond to changes in the intensity of the DAF signal can be used to determine whether speakers treat DAF as speech or noise (Howell, 1990). When speakers heard their non-delayed speech amplified whilst speaking, voice level decreased (Howell, 1990) which indicated speakers compensated (Fletcher effect) intensity when vocal output was too loud (Lane & Tranel, 1971). However, when the intensity of speech delayed by 100ms was increased, speakers *increase*d the intensity of their speech. (Howell, 1990) This is a Lombard effect that occurs when speech is produced in noisy environments (speakers attempt to speak above the noise). The Lombard effect indicated that the delay transformed speech into a non-speech noise, rendering it unusable for feedback control. The Lombard effect in Howell (1990) was monotonic over the intensity range studied (all increases in DAF intensity increased vocal intensity). A Lombard effect also occurred when speaking along with white noise which further confirmed that the delayed speech sound under DAF was treated as noise rather than speech (Howell, 1990).

For fluent speakers, the length of the DAF delay-interval affects speech rate (Howell & Powell, 1987). The effects on speech rate differ for “short” delays (<100ms) where speech rate reduces, and “long” delays (>100ms) where speech rate increases (Kalinowski et al., 1996). Note that other authors classify delays as short/long at different durations (discussed below).

The length of the delay interval and/or intensity of DAF have been manipulated with PWS. In general, choice of what parameter value to use depend on effects Speech and Language Pathologists (SLPs) might want to induce when using them in interventions. For instance, DAF-delays could be chosen that slow speech rate or the level of DAF could be adjusted to increase vocal intensity.

[1] Speech error here includes medial vowel elongations. If medial elongations are not considered to be an error then the estimated effect of DAF would be less.

## People who stutter

Lee (1950) considered that there were similarities between the dysfluencies induced in fluent speakers when speaking under DAF and the dysfluencies that occur under normal-speaking conditions in PWS. The supposed similarity is misleading (Howell, 2011). For instance, superficially prolongations in stuttered speech are similar to the vowel-elongations noted in connection with DAF. However, prolongations usually occur on onset consonants whereas medial vowels are mainly affected under DAF (Howell et al., 1988). Thus, prolongations and vowel-elongations under DAF differ both on phone type, and syllable position that is affected. Nor is the speech of PWS like that observed when fluent speakers receive DAF: The natural speech of PWS is not louder or more monotonic as occurs when fluent speakers are given DAF (Howell, 2011).

Nevertheless, DAF improves the fluency of PWS to some extent. Early research on DAF with PWS tended to use long delays (100ms and above) and concluded that DAF has the robust effect of removing stutters (Goldiamond, 1965; Lotzmann, 1961; Nessel, 1958; Soderberg, 1969). That said, presenting long-delay DAF to PWS has undesirable side effects: The vowel-elongation noted to occur when fluent speakers speak under DAF also occurs with PWS (Howell et al., 1988). Howell (1990) reported that vocal intensity of PWS under long-delay (100ms) led to a Lombard effect as noted with fluent speakers. This suggests that PWS process DAF at this delay as noise similar to fluent speakers. The increased vocal intensity under long-delay DAF that PWS experience, causes DAF speech to sound unnatural and the pattern persists post-intervention (Howell, 2004a; Novak, 1978). As well as external adjustments to intensity, speakers can influence feedback level by speaking at different loudness levels. If speech is at a low level, fluency-enhancement does not occur (Butler & Galloway, 1957). The intensity level of the delayed speech is important in accounting for differences in susceptibility to DAF across speakers, considered below (Howell & Archer, 1984).

Speech rate, vocal intensity and stuttering rate of PWS depend on the delay selected (Howell, 1990; Lotzmann, 1961; Soderberg, 1969). In the early studies on PWS, delays were employed that produced maximum disruption to fluent speakers (100-200ms.). Although long-delay DAF-speech reduces stuttering symptoms, the ancillary disfluencies that occur (drawls, loud speech, flat pitch and slow speech) are noticed by listeners (Geetha et al., 2017; Stuart & Kalinowski, 2004) and, as noted, may persist in post-intervention speech (Novak, 1978). This has implications about how to employ DAF in interventions for PWS as one type of speech dysfluency may be replaced with another form.

The finding that the effects of DAF differ between long, and short delays with fluent speakers, is supported by studies on PWS. Kalinowski et al. (1996) proposed that DAF-delays less than 100ms were ‘short’ and that delays of 75ms and 50ms were optimal for improved speech rate and fluency. In unpublished work, Howell systematically varied delays below 100ms and observed the impact on vocal output intensity when feedback intensity was adjusted at each delay. Delays under 30ms led to a Fletcher effect whilst longer delays led to a Lombard effect. This suggested that the delayed sound was only treated as speech at delays up to 30ms (somewhat shorter than Kalinowski et al. 1996 proposed).

Contemporary prostheses and research equipment delivers short-delay DAF (Howell, 2004a), typically 20-60ms (Kalinowski et al., 1993). Short-delay DAF improves speech naturalness compared to long-delay DAF (Kalinowski et al., 1996) and maintains the fluency-enhancing effects observed under long-delay DAF (Antipova et al., 2008; Kalinowski et al., 1993; Kalinowski et al., 1996; Sparks et al., 2002; Stuart et al., 2004; Van Borsel et al., 2003). The improved naturalness probably arises because short-delay DAF (Kalinowski et al., 1996), like synchronous spectrally-changed speech (Howell et al., 1987), does not affect speech control as described for long-delay DAF. The findings that certain forms of AFB, including short-delay DAF, improve fluency when there is no global change in speech rate suggests that slowing speech is not the operative mechanism behind improvements in fluency of PWS. It has been claimed that short-delay DAF produces effects on speech fluency, rate and naturalness that are as good as other forms of altered sensory feedback (Kalinowski, et al., 1993; Macleod, et al., 1995) although this has been contested (Howell & Sackin, 2000, 2002; Kalveram, 2001; Kalveram & Jäncke, 1989).

## Clinical work using DAF with PWS

Short-delay DAF would be preferred if this form has equivalent effects to long-delay DAF in terms of stuttering reduction but does not slow speech rate nor increase vocal intensity. In the light of this, it is surprising that Van Borsel et al. (2003) found that PWS preferred delays of 93-147ms when using a DAF-prosthesis over a 3-month period. Soderberg (1969) also reported PWS preferred long-delay DAF whereas Kalinowski et al. (1996) and Lotzmann (1961) reported a preference for short-delay DAF. We recommend that delay length should be optimised for PWS by allowing them to adjust delay between 0 and 250ms.

The fluency-enhancing effects of DAF occur predominantly whilst stimulation is delivered (Lincoln et al., 2006) which confirms its prosthetic role. This allows users to choose when to use their prostheses (e.g. at interviews or when speaking in public). Long-delay DAF can also be used as a way of reducing speech rate in multi-component interventions (Atkinson-Clement et al., 2015; Ryan, 1974). One such intervention that uses DAF (Basi et al., 2016; Bothe et al., 2006) is Ryan's (1974) treatment program where the DAF component is based on Goldiamond's (1965) work.

Moreover, ways of promoting carry-over with DAF and other AFB-prostheses merit further investigation (Howell, 2004a). PWS do not need AFB altered all of the time since stuttering is intermittent. Hence, targeting altered AFB on sections of speech where there are stuttering symptoms may be sufficient to reduce dysfluency as the fluency-enhancing effects happen immediately (Howell, 2004b). Another alternative would be to reduce dosage by ramping intensity of altered AFB down after a dysfluency provided speech is fluent and switch it to its full level when the next episode of stuttering occurs (leaky integrator). Additionally, presenting altered AFB on every stuttering symptom might not be optimal for promoting long-term fluency (Howell, 2004b). The Partial Resistance to Extinction Effect or 'PREE' (Hochman & Erev, 2013) indicates better retention of behaviours when reinforcement is given intermittently according to a specified schedule rather than on every occurrence. Therefore, with respect to PWS, it may be preferable to deliver fluency-inducing AFB only on a proportion of moments of stuttering that could promote retention of altered AFB's fluency-enhancing effects to unaided speech. This could be implemented if a prosthesis was controlled by an automatic recognizer that identifies stuttering symptoms (Howell et al., 1997a & Howell et al., 1997b) which is pre-set to deliver altered AFB on a proportion of the dysfluencies detected. Reed and Howell (2000) proposed a framework for implementing such procedures, but empirical work needs to be conducted. The present situation is that altered AFB is usually presented continuously whilst clients speak, which according to the above analysis may not be the best way of promoting retention of fluent behaviour.

A potentially more serious issue than fluency improvements only occurring whilst DAF is delivered, is the claim that speakers 'adapt' (lose sensitivity) to DAF and other AFB manipulations. Adaptation has similarities with the remission of stuttering noted in connection with other treatments (Weidig, 2005). The supposed adaptation under AFB would benefit from examining work on why fluent participants show differences in susceptibility to

DAF (Howell & Archer, 1984). It is assumed that the reason for fluctuations that occur across individuals (susceptibility) are related to those that happen within an individual across occasions (supposed adaptation in some PWS). Intra-individual differences in susceptibility to DAF in fluent speakers are due to adjustments of own-voice level (Howell, 2004b). Put simply, if speakers silently-mouthed utterances, there would be no auditory signal to delay and no effect of the manipulation, whereas speakers who try to shout over the DAF, enhance the altered sound. PWS using prostheses long-term may subconsciously adapt voice intensity to lower levels to make speech easier to produce that also reduces DAF-dosage. Automatic gain-control circuits (AGCs) can be included in devices to control for variations in voice levels over occasions (Howell, 2004a). AGCs ensure AFB is at a constant level irrespective of adjustments to voice level by speakers. AGCs could be included in clinical devices as they should produce a more uniform response across speakers and time according to this analysis.

Overall, temporal delays to speech feedback have a robust effect on fluent speakers and PWS. In fluent speakers, DAF induces a range of disruptions to speech (makes it sound dysfluent in different ways to stuttered speech). Two important concepts that apply to both speaker groups are: (1) short-delay DAF produces no Lombard effect, and (2) long-delay DAF (>100ms) is responded to as noise (speakers increase vocal intensity. The delay length may, therefore, determine how delayed feedback is processed by the speaker; with longer delays being regarded as noise that affects the sensorimotor system (increased voice level) whereas, shorter delays retain processing as speech. The effects of DAF and how these depend on delay-length and participant group require further examination. DAF provides a tool to alter certain aspects of speech production for researchers and SLPSs. However, differences in susceptibility need further examination as there are interesting possibilities concerning how to maximize benefit in prostheses (AGCs). A further thing to examine with prostheses, is control of DAF dosage (PREE) using procedures that incorporate automatic recognition of stutters.

## Frequency altered feedback

We need to review how speech is generated to understand procedures that alter frequency components of speech. Voiced speech results in a harmonic sound complex that gives speech its pitch. The pitch percept of this sound is dominated by the lowest harmonic (the fundamental, F0). The harmonic complex also excites the resonant frequencies of the vocal tract, called formants (Rosen & Howell, 2011). The formants are numbered from the lowest resonance upwards. The low-order formants (first and second, F1 and F2) are most important for speech intelligibility.

Some, or all parts of the speech spectrum can be shifted in real time using a variety of software and hardware devices. Variants include shifts of the whole spectrum (frequency shifted feedback, FSF[2]; Elman, 1981; Howell et al., 1987), F0, F1 or F2 separately (Donath, Natke & Kalveram, 2002; Houde & Jordan, 1998, 2002; Jones & Munhall, 2000, 2002; Natke & Kalveram, 2001; Purcell & Munhall, 2006a & Purcell & Munhall, 2006b). F0, F1 or F2 have to be extracted from speech before they can component alone can be selectively shifted. DSP chips perform the computations rapidly but methods that rely on the Fast Fourier Transform (FFT) require a sample before they can start the computation. Waiting for the sample before computation takes place delays AFB (AFB would then be shifted component plus DAF). Speed-changing methods that sample amplitude fluctuations in signals digitally at a rapid rate and replay them almost instantaneously at a changed rate do not entail delays (e.g. Howell et al.'s 1987 FSF procedure). The work reviewed in this section looks at shifts of: (1) whole spectrum (FSF); (2) F0; and (3) single formants.

All types of shifted sounds can be delivered in three paradigms which address distinct questions and require different response measures: (1) sustained shift during speech (often used for eliciting fluency improvements in patient groups); (2) immediate *compensation* where an attribute of speech is shifted unpredictably for a short time (typically 500ms). Compensation occurs when the speaker responds to offset the change; (3) *adaptive* presentation where, after baseline response is measured, a component of speech (F0, F1, F2 etc.) is gradually shifted, the shift is sustained for a period before it is gradually changed back to baseline ,which is rechecked (Purcell & Munhall, 2006a). As well as any compensation that occurs in the shift phase, the changed response may persist beyond the shift period (does not return to baseline) indicating an adaptive response to the altered component (shows that a novel speech-motor command has been learned).

Finally, the speech component manipulated in the procedures and the responses measured do not always correspond. For instance, Elman (1981) shifted the whole speech spectrum but only looked at the impact this had on F0. Other possibilities (e.g. effects on pitch responses when formants are shifted) remain to be investigated in fluent speakers as

[2] Note, Frequency Shifted Feedback or FSF refers to the shifting of the whole spectrum, not specific manipulations like F1 perturbation.

well as speakers with speech disorders. Procedure section headings refer to the response measured.

## Fluent speakers

### Whole spectrum and F0 perturbation

Elman (1981) investigated what impact shifting the entire speech spectrum had on F0 control. The incoming speech signal was compressed in real time during speech, which shifted it upwards in frequency by approximately 10%. The perturbation induced an almost immediate opposing shift (F0 was lowered).

Such compensatory responses of F0 in fluent speakers have been replicated (Burnett et al., 1998; Chen et al., 2007; Liu & Larson, 2007). However, whilst most participants compensate (change pitch in the opposite direction to the imposed shift), some follow the change in pitch and a few show no change (Burnett et al., 1998). Such inter-individual differences have been attributed to reliance either on auditory or somatosensory feedback (Lametti et al., 2012). Whilst this proposal may account for cases where compensations or null responses occur to frequency-shifted AFB, it does not explain why a significant minority of speakers follow the shift. The latter appear to rely on auditory feedback (they respond), but the response occurs in a non-compensatory way.

No Lombard or Fletcher effect occurs when the whole speech spectrum is shifted for fluent speakers (Howell, 1990). Thus, unlike long-delay DAF, intensity is not affected by FSF.

Although FSF does not affect voice-level responses, it does affect F0 (Elman, 1981) and, hence, respiratory, laryngeal, and articulatory control. Thus, Heinks-Maldonado and Houde (2005) reported a relationship between voice-level perturbation and F0: a rapid 400ms shift in voice-level (±10dB) whilst phonating a vowel resulted in speakers compensating by decreasing or increasing the F0 of their speech output. Subsequent studies provided more detail about this response: Bauer et al. (2006) highlighted that the magnitude of the F0 response diminished as the voice-level perturbation decreased (±1dB resulted in smaller F0 compensations compared to ±3dB or ±6dB). Larson et al. (2007) investigated whether the compensatory responses to F0 shifts were influenced by shifts in vocal loudness. In three experiments, F0 frequency was shifted: (1) upwards or downwards during phonation; (2) vocal intensity level was increased or decreased by 3dB; and (3) the two perturbations were combined. The compensatory responses for the shifts in decreasing order were 18cents when

F0 alone was altered, 14-16 cents for simultaneous shifts of F0 and vocal intensity and 10 cents when vocal intensity alone was altered. Spectral-compensation magnitude also depended on the frequency of F0 where higher F0 values led to greater spectral-compensatory shifts and reduced response times to shifted speech (Liu & Larsson, 2007). Downward F0 shifts elicit larger compensatory responses than upward shifts (Chen et al., 2007). It is unclear why direction of perturbation elicits different responses. However, these findings can be exploited in speakers with fluency problems when adjustment of F0 is required in interventions. The F0 magnitude (Larsson et al., 2007) and direction response effects (Chen et al., 2007) decay rapidly after FSF is switched off (F0 returns to baseline quickly). However, FSF can have long-term effects if the perturbation is sustained for long periods of time (Munhall 2006a). Hence, procedures to promote carry-over (discussed in the section on DAF) are worth exploring.

Perturbations to F0 occur with the following materials: sustained vowels (Bauer & Larson, 2003; Hain et al., 2000; Larson et al., 2001; Sivasankar et al., 2005), glissandi (Burnett & Larson, 2002), song (Natke et al., 2003), nonsense syllables (Donath et al., 2002; Natke et al., 2003; Natke and Kalveram, 2001) and Mandarin phrases (Jones & Munhall, 2002). Perturbation of F0 also changes the supra-segmental timing of utterances (Bauer, 2004). Furthermore, shifting feedback on one syllable affects the output of the next one (a coarticulation effect). Thus, shifting F0 either upwards or downwards resulted in an upward shift in F0 on the subsequent syllable (Natke & Kalveram, 2001). Similar effects occur in immediate compensation paradigms where transient shifts to F0 resulted in changes in the voice fundamental on subsequent material (Donath et al., 2002).

### Formant perturbation

F1 frequencies are higher for open, than for close, vowels. F2 frequencies are higher for front, than back, vowels. Similar relationships apply with consonants but the picture is more complex because articulatory shape changes dynamically during their production. Perturbation of F1 or F2 could imply that articulation of speech is incorrect. Therefore, in an adaptation paradigm, participants should adjust to compensate for the formant perturbation (Purcell & Munhall, 2006a; Villacorta et al., 2007).

Adaptation and compensation paradigms have been used to study F1 and F2 perturbations in fluent speakers. In both cases, most speakers adjust their formant frequencies in the opposite direction to the perturbation. Again (as with FSF and F0 perturbations), there are some individuals who follow and others who do not respond (Purcell & Munhall, 2006a).

When F1 or F2 are manipulated, the magnitude of the response is proportional to the magnitude of the perturbation providing the perturbation exceeds approximately 60Hz (Purcell & Munhall, 2006a). Specifically, larger perturbations of F1 and/or F2 result in greater changes to the corresponding F1 and/or F2 values in the resultant speech (Cai et al., 2011; Purcell & Munhall, 2006b). Unlike FSF and F0 shifts however, perturbations of F1 and F2 do not result in responses larger than about 15% of the imposed shift (Purcell & Munhall, 2006b). For F1 changes, as with F0 perturbations, downward shifts elicit larger responses than upward shifts at 16.3% and 10.6% of the applied formant shift, respectively (Purcell & Munhall, 2006b). Cai et al. (2011) reported how F1 and F2 manipulation affect their corresponding responses but did not look at non-corresponding responses.

There is no evidence whether perturbation of F1 or F2 leads speakers to alter their vocal intensity (i.e. produce a Lombard, or Fletcher, effect). It is worth checking whether speakers who compensate produce a Fletcher effect whereas those who do not compensate produce a Lombard effect (treat the shifted speech as speech, or noise, respectively). Also, it is not known whether varying intensity of F1 or F2 selectively affects the intensity of the corresponding formant or the overall intensity.

In sum, fluent speakers show both compensation and adaptation to manipulations of F1 and/or F2. This reflects both the malleability of speech-sound representations at a neuro-motor level and the ability to accurately adjust current speech output in response to concurrent auditory information in fluent speakers. It is interesting to consider why some speakers respond to the frequency perturbation by opposing the shift, whilst others follow and yet others do nothing whatsoever. All these responses are likely to arise at low levels of neural function as people are not usually aware of frequency perturbations nor their own compensation or adaptation response.

## People who stutter

### Whole Spectrum and F0 perturbation

Howell et al. (1987) shifted the whole speech spectrum of PWS and reported a marked reduction in concurrent stutters. This reduction under FSF has been replicated using monologue, conversation and reading materials and for a range of frequency shifts usually up to an octave (Armson & Stuart, 1998; Hargraves et al., 1994; Kalinowski et al., 1993; Natke et al., 2001; Ritto et al., 2016;  Stuart & Kalinowski, 2004). Fluency enhancement does not correlate with the magnitude of the shift and there is no particular shift that has maximum effect. For instance, Stuart et al. (1996) suggested that the entire spectrum had to be shifted by at least ¼ of an octave to induce fluency-enhancing effects in PWS but upward shifts of ¼ octave and above all led to similar fluency enhancements for PWS (Kalinowski et al., 1993; Hargrave et al., 1994; Stuart et al., 1996). Some studies report that groups of PWS only compensate to upward FSF perturbations (Bosshardt et al., 1997; Natke et al., 2001). However, downward shifts are known to reduce the likelihood of F0 compensations (Natke et al., 2001) and to promote fluency (Bauer et al. 2007; Natke et al., 2001). As with fluent speakers, it is not known why direction-specific effects occur and, additionally for PWS, whether these indicate differences in underlying neural pathology. Nevertheless, this review suggests that downward shifts are recommended for clinical use, assuming F0 compensations should be avoided, even though upward shifts are studied most often.

FSF improves speech fluency without adversely affecting other aspects of speech including intensity (Howell, 1990): FSF does not affect speech naturalness in PWS other than partial compensatory shifts in F0 (Natke et al., 2001); FSF does not slow speech (Howell, et al., 1984; Howell & Sackin, 2000; Natke et al., 2001).

There is no work that has studied individual susceptibility to the fluency-enhancing effects of FSF *a priori* (studies have looked at individual differences *post facto*) Although PWS differ in whether or not they respond to F0 perturbations, compensatory response to F0 perturbation does not predict fluency-enhancing responses to FSF (Natke et al., 2001). The fluency enhancement does, however, seem to relate to the implementation method with speech-change procedures proving most effective (Howell, 2004a). Although the fluency-enhancing effects of short-delay DAF have been reported to be 'equal' (i.e. not significantly different) to that of whole-spectrum FSF, DSP implementation of FSF were used in the studies (Hargreaves et al., 1994; Kalinowski et al., 1993; Stuart et al., 1994, 2004, 2006). As mentioned, these introduce delays because they use the FFT to shift the speech spectrum (Howell & Sackin, 2000). Thus, estimating an FSF effect separate from short-delay DAF is not possible in these studies (Howell & Sackin, 2000).

## Clinical trials using FSF prostheses

The reasons FSF is commended for clinical work are that it has few side effects on intensity, speech rate and F0, but large effects on fluency. These findings led to clinical trials to establish whether enhanced fluency under FSF persists over time. Stuttering symptoms decrease by up to 80% in PWS under FSF (Armson & Stuart, 1998; Howell et al., 1987; Howell 1990; Kalinowski et al., 1993; Natke et al., 2001). Consequently, prostheses that employ FSF have appeared including SpeechEasy© (Janus Development Group Inc.), Casa Futura Technologies, VoiceAmp, National Association for Speech Fluency, A.S. General Limited, Digital Recordings and Kay Elemetrics. There are important differences in the way FSF is implemented in these prostheses and as compared to laboratory set-ups. Prostheses that are worn have limited processing power. Consequently, prostheses often have significant additional delays between the speech input and when the perturbed output is returned to the speaker. Unlike laboratory implementations, SpeechEasy© loses harmonic information when frequencies are shifted as it applies an absolute shift to the entire speech signal (Stuart et al., 2003). For example, if a harmonic complex of 500, 1000 and 2000Hz (first, second and fourth harmonics of 500 Hz) received a 400Hz shift, components would be at 900, 1400 and 2400Hz. The latter components are not harmonically-related which would affect perceived pitch of the complex. Delays and distortions to harmonics are important when considering the efficacy of prostheses.

FSF prostheses improve fluency in PWS and provide an effective treatment for stuttering (Armson et al., 2006; Kalinowski et al., 1993; Stuart et al., 2003, 2004). However, the effect size is not as large as laboratory results would suggest whether comparisons are made with the sampling or DSP methods. Thus, Armson et al. (2006) documented the effect sizes for the SpeechEasy© device in different speaking tasks both in and out of their laboratory. Reductions in laboratory conditions were about 49% for conversations, 36% for monologues and 74% for readings (Armson et al., 2006). When the SpeechEasy© was used outside the laboratory, no reduction was apparent either in “Situation of Daily Living” (SDL) or at follow-up (O’Donnell et al., 2008). Rate of stuttering even increased for some participants when the device was used for 16 weeks.

Gallop and Runyan (2012) reported similar findings over a longer time-period (59 months). Reductions in stuttering for the whole group did not occur but some individuals showed appreciable reductions after 59months of use (others showed increased stuttering rate). Extended use also changed the stuttering frequency from before the prosthesis was used, even when the prosthesis was not used frequently by individuals. The reasons for these individual differences are not clear. Assuming, as argued with fluent speakers, that changes to speech control are mediated by neural changes, participants may have either adapted to maintain the fluency-inducing effect or to ignore the perturbation and continue speech as they did before baseline. The findings highlight that brains respond to long-term perturbation in different ways. Intermittent presentation of FSF could prevent extinction of responses (PREE described earlier). Consistent with this, although most applications have applied FSF throughout speech, it has been focussed on syllables onsets and this 'pulsed' form of FSF provides the same level of fluency enhancement as continuous FSF (Howell et al., 1987). Implementation of pulsed FSF in prostheses would require increased processing power which may explain why continuous presentation is used exclusively in current devices.

Despite prostheses providing a low-cost alternative to conventional treatments intended to reduce stuttering, the long-term effects remain unclear and may even be detrimental to certain individuals. Theoretically-motivated ways of reducing adaptation remain to be explored (Howell, 2004b). Furthermore, from a conventional perspective, there is no consensus about what factors could predict whether an individual will respond positively or negatively, although, some research has suggested that PWS with a mild stutter are more likely to be affected detrimentally (Gallop & Runyan, 2012; O'Donnell et al., 2008). Nevertheless, FSF provides researchers and clinicians with a technique to induce fluency, at least temporarily, in most PWS.

It is interesting to consider how a speaker adjusts to these perturbations over a long period of time from a neuroplasticity perspective. As detailed, the extended use of AFB-prostheses generally results in a diminished fluency-enhancing effect (O'Donnell et al., 2008). Findings from all studies on PWS that observed effects over extended time-periods report highly variable response (Gallop & Runyan, 2012; O'Donnell et al., 2008; Ritto et al., 2016). Note that neither the longitudinal studies (Gallop & Runyan, 2012; O'Donnell et al., 2008) nor the randomised control trial (RCT) of Ritto et al., (2016) included a non-treatment control group of PWS for comparison. However, from research on formant manipulations detailed below, the brain adapts and recodes speech processes in response to manipulated feedback signals (Purcell & Munhall, 2006a). Although this research is restricted to formant manipulations, it is possible that extended use of whole-spectrum FSF causes an adaptive response from the brain. Unfortunately, no research to date has examined the neurophysiological effects of extended presentation of FSF to assess this point. Further research needs to elucidate whether neuroplastic changes to protracted AFB: (1) occur, and (2) whether the changes aid long-term fluency versus whether they lead to dependency on the prosthesis.

In sum, due to technological constraints, prostheses induce fluency improvement in PWS but not to the extent reported in lab studies. Nonetheless, the improvements are significant and are at least as good as current behavioural therapies at a group level (Ritto et al., 2016). As discussed, responses to the prostheses vary markedly in different contexts (Armson et al., 2006) and the outcomes after long-term use are not clear. Again, it must be stressed that prostheses: (1) include delays (both long-delay and short-delay) in conjunction with FSF adding a further factor that makes responses variable; and (2) perturb entire utterances which is not necessary and possibly disadvantageous insofar as it resulted in several users criticising the accompanying noise (Pollard et al., 2009).

### F0 perturbation

Scanning evidence suggests that PWS have a deficit in sensorimotor integration (Guenther & Tourville 2011; Watkins et al., 2008). This has been tested in tasks where responses to F0 perturbations have been examined, Consistent with the proposed deficit, PWS show a variable and diminished compensatory response to the instantaneous F0 perturbation compared to fluent speakers (Sares et al., 2018).

PWS show similar responses to fluent speakers when magnitude of spectral perturbation to F0 or speech intensity are varied, but responses are reduced (amplitude) and delayed (phase) for PWS (Loucks et al., 2012). The deficient responses indicate an inability for PWS to appropriately account for perceived errors both in strength and timing of corrective response, compared to fluent speakers. Bauer et al. (2007) reported that the abnormal responses in both amplitude and phase in PWS did not apply when the magnitude of the shift was 600 cents. It is unclear why PWS, compared to fluent speakers, are no less able to respond to F0 perturbations when the magnitude of the perturbation is relatively large but, unable to respond adequately to smaller perturbations (100cents in Loucks et al., 2012). Unlike FSF, F0 perturbations have no reported clinical benefit and do not promote fluent speech.

### Formant perturbation

As noted previously, F1 and F2 perturbations provide information about how an individual integrates, compensates and adapts their articulator movements. PWS show similar responses to fluent speakers in: (1) adaptation where, after formants are incrementally shifted over utterances, PWS reconfigure articulatory actions to restore the perceived (afferent) speech to match more closely to the intended speech (efferent) output (Daliri et al., 2018); and (2) immediate compensation where PWS compensate for unexpected F1 and F2 perturbations by adjusting their articulators by reducing or increasing formant values (Cai et al., 2012, 2014).

There are, however, some subtle differences between PWS and fluent speakers: Unlike response to F0 where PWS showed lagged response times to F0 perturbation (Loucks et al., 2012), PWS took a similar time (150ms) to fluent speakers to adjust to formant perturbation (Cai et al., 2012). The magnitude of compensation was significantly reduced compared to fluent speakers (Cai et al., 2014). These findings suggest a deficit in ability to integrate spectral, but not temporal, formant information. Furthermore, the differences noted between PWS and fluent controls only apply to adult PWS since children's responses were not distinguishable from their fluent peers (Daliri et al., 2018). This suggests that adults acquire this deficit.

There is no evidence whether formant perturbations influence vocal intensity. PWS may have a similar response to the one they have to FSF (intensity reduces slightly, Howell, 1990). Also, although there is no research about whether formant perturbation affects speech rate, it seems likely that it would have a similar effect to FSF and F0 shifts (no slowing of speech rate).

Although formant manipulation permits better probing of what aspects of speech are deficient in PWS, there are no known clinical benefits in using such forms of altered feedback as yet. Nevertheless, since whole-spectrum shifts (which includes the formants) improve speech fluency, formant manipulations may help reveal the as-yet unknown mechanisms driving this fluency enhancement.

## Vibrotactile feedback

Several parts of the body vibrate during articulation and provide sensory information about speech activity. As with altered *auditory* feedback, vibration can be altered to investigate how this form of afferent input affects speakers. Vibration during articulation possibly only carries information about F0 (Howell & Powell, 1984). Nevertheless, there are several features of vibrotactile feedback (VTF) that make it interesting. VTF: (1) operates over different neural pathways to AFB (Cheadle et al., 2018); (2) does not interfere with auditory information[3];

[3] Note, VTF can generate auditory signals if presented at a too high a frequency and intensity. Low resonant frequencies of ~100Hz are unlikely to generate an auditory signal. However, 230-250Hz are often used in the literature as this allows maximal stimulation of Pacinian corpuscles (Griffin, 1990; Siegel & Sapru, 2006) and stimulating at this high frequency may provide some undesired noise and therefore masking.

and (3) induces corresponding fluency changes to AFB (Cheadle et al., 2018; Kuniszyk-Jozkowiak et al., 1996; Snyder et al., 2009; Waddell et al., 2012).

Vibratory information is sensed by two types of mechanoreceptors in the skin which code for the intensity and frequency of the vibration: Meissner's and Pacinian corpuscles (Figure 1). Meissner's corpuscles code for 20-70Hz vibrations and their resonant frequency is 30Hz (Iggo et al., 1986). Pacinian corpuscles code for 100-400Hz vibrations and have a resonant frequency of 250Hz (Griffin, 1990; Siegel & Sapru, 2006). Vibratory stimulation at different frequencies, therefore, activates different mechanoreceptors. Vibratory information ascends to the primary somatosensory cortex via two main pathways: (1.) the main sensory trigeminal cranial nerve (CN), transducing stimulation to the facial and scalp areas including the vocal system, and (2) the posterior column-medial lemniscal (non-CN) pathways, carrying all stimulatory information presented to everything caudal to the cranium (Juliano & Mclaughlin, 1999). In particular, vibrotactile stimulation at the facial and cranial tissues will ascend through the CN-pathway whilst stimulation at all parts of the body otherwise ascend through the non-CN pathway. The CN-pathway also transmits vibratory information through the brainstem at the level of the mid-pons, the trigeminal lemniscus and the thalamus. The non-CN pathway transmits to the medulla, thalamus and primary somatosensory cortex. Note that both pathways meet at the thalamus and, unlike AFB, neither carries information to the cerebellum (Hendelman, 2005; Juliano & Mclaughlin, 1999).

## Fluent speakers

Work on VTF with fluent speakers has mainly been conducted on perception rather than production. Thus, Fucci et al. (1985, 1991) compared effects of VTF on fluent speakers (and PWS) in perceptual work. They investigated individual sensitivity to vibrotactile stimulation at 250Hz (activated the Pacinian corpuscles), delivered to the hand and the tongue. The experimenter controlled a master attenuator and participants had their own attenuator to modulate the intensity of stimulation. Intensity was gradually increased to the threshold of sensation for the participant. The experimenter then modulated the intensity of stimulation and the participants had to follow this modulation by adjusting their own attenuator. In effect, they estimated how the participant judged that the experimenter had changed intensity. Fluent speakers had more accurate representations of somatosensory information emanating from the tongue, not the hand, compared to PWS. PWS integrated somatosensory information from the arm as well as fluent speakers (Fucci et al., 1985). This suggests that VTF should have little or no effect on speech in fluent speakers as their somatosensory feedback for speech is good. Further work using VTF with fluent speakers is required to verify this.

With respect to production studies, VTF has not usually included a fluent control group (Cheadle et al., 2018; Kuniszyk-Jozkowiaket al., 1996, 1997; Synder et al., 2009; Waddell et al., 2012). Kuniszyk-Józkowiak and Adamczyk (1989) is the only study on the effects of VTF on fluent speakers and results were reported as comparisons with PWS. This is discussed in the following section.

## People who stutter

Fucci et al.'s (1985, 1991) investigations tested whether PWS have poor sensorimotor integration for vibratory information similar to that claimed for auditory stimulation (Cai et al., 2014; Loucks, Chon & Han, 2012). Their perceptual studies, discussed in the previous section, showed that fluent speakers have more accurate somatosensory representations about the articulators relative to PWS.

Research has considered how VTF affects fluency of PWS and how this varies when VTF is perturbed. VTF improves fluency (Cheadle et al., 2018; Kuniszyk-Jozkowiaket al., 1996; Synder et al., 2009; Waddell et al., 2012).

The intensity of VTF affects speech production in PWS, however the effect shows a quadratic relationship with intensity: Low (0.5–1 $m/s^2$) and high (2.5–3 $m/s^2$) amplitudes of vibration induce greater reductions in stuttering than mid-range (1–2.5 $m/s^2$) amplitudes (Cheadle et al., 2018). Placement of the vibrator, and therefore the neural pathway stimulated, did not influence stuttering frequency during VTF (Cheadle et al., 2018). This contrasts with the perceptual work that found vibration applied to the articulators had different effects than when vibration was applied to the hand for PWS compared to fluent speakers (Fucci et al., 1985).

Kuniszyk-Józkowiak and Adamczyk (1989) included PWS and compared their results with a control group. Concurrent and delayed VTF were compared for stimulation delivered separately to the left and right middle fingers at 230 Hz with 0.5mm maximum amplitude of oscillation. There were no differences in speech rate in response to VTF between fluent speakers and PWS rather, speech rate slowed for all participants when stimulation was delayed. The rate-slowing effect is at odds with Cheadle et al. (2019) who found that, PWS did not change their speech rate in response to VTF whether the stimulation was concurrent or delayed. However, procedural differences may account for this discrepancy, Cheadle et al.

stimulated at a lower peak frequency (~100Hz) whereas, Kuniszyk-Józkowiak and Adamczyk (1989) used a high resonant frequency (230Hz). Hence, there are two possible accounts for the difference:

1. Compared to Cheadle et al., Kuniszyk-Józkowiak and Adamczyk stimulated at a higher resonant frequency than would activate Pacinian corpuscles. Thus they may have been more successful then Cheadle et al. at perturbing speech when VTF was delayed and this led to a reduction in speech rate.
2. High frequency VTF stimulation (Kuniszyk-Józkowiak & Adamczyk, 1989) can result in audible noise. The noise arising from the delayed VTF would act like DAF when delayed and would slow speech (Howell et al., 1992).

As mentioned, the neural pathway for VTF differs from the auditory pathway until the post-thalamic pathways. Consequently, VTF does not pass through the cerebellum and therefore, VTF cannot affect timekeeping processes in the lateral cerebellum reasoned to be implicated in fluent speech control (Howell, 2011). This may explain why DAF has greater effects on stuttering reduction (up to 85% according to Lincoln et al., 2006)) compared to VTF (maximum of 80%). Furthermore, Cheadle et al.'s (2018) recent study on VTF with PWS found effect size to be 21.82% across all conditions. Finally, the way the effect size varies as a function of intensity, frequency of vibration, frequency of use, timing of vibration and inter-individual differences are not fully understood. This point is underlined by the lack of such information on other patient groups and fluent speakers.

Although early research showed that delaying VTF further reduced dysfluency (Kuniszyk-Jozkowiak et al., 1996), this was not replicated by Cheadle et al. (2018). Hence, the role this parameter could play in clinical applications using VTF is unclear. No research has looked at whether auditory parameters of speech such as, F0 and formant frequencies change with respect to VTF for PWS.

# Concluding Remarks

Speech production can be modulated by altering sensory feedback. With respect to PWS, all procedures that have been reviewed promote fluency. The principles and parameters of each sensory feedback manipulation have been outlined and areas where future work is needed have been indicated. Clearly the effects of altered sensory feedback are diverse and vary by technique, whether individuals are fluent or not (PWS) and between individuals within each fluency group.

Looking at fluency-enhancing effects in PWS, the techniques lend themselves, and have been implemented, as prostheses. As the fluency-enhancing effects depend on the prosthesis being active, there should not necessarily be any carry-over of fluency enhancement. However, the alterations induced by long-term use of the devices may lead to neuroplastic changes that change speech permanently (Purcell & Munhall, 2006a). Present implementations of techniques to alter sensory feedback in prostheses are not optimised to counteract adaptation and to achieve effects at low dosages. We have suggested prosthetic procedures may need to be modified so that they are only active for a percentage of utterances on parts of the utterance (Hochman & Erev, 2013; Reed & Howell, 2000)) at points where problems are being experienced (Howell & El-Yaniv, 1987).

# References


Antipova, E. A., Purdy, S. C., Blakeley, M., & Williams, S. (2008). Effects of altered auditory feedback (AAF) on stuttering frequency during monologue speech production. *Journal of Fluency Disorders*, *33*(4), 274–290. https://doi.org/10.1016/j.jfludis.2008.09.002

Armson, J., Kiefte, M., Mason, J., & De Croos, D. (2006). The effect of SpeechEasy on stuttering frequency in laboratory conditions. *Journal of Fluency Disorders*, *31*(2), 137–152. https://doi.org/10.1016/j.jfludis.2006.04.004

Armson, J., & Stuart, A. (1998). Effect of extended exposure to frequency-altered feedback on stuttering during reading and monologue. *Journal of Speech, Language, and Hearing Research*, *41*(3), 479–490. https://doi.org/10.1044/jslhr.4103.479

Atkinson-Clement, C., Sadat, J., & Pinto, S. (2015). Behavioral treatments for speech in Parkinson's disease: meta-analyses and review of the literature. *Neurodegenerative Disease Management*, *5*(3), 233–248. https://doi.org/10.2217/nmt.15.16

Basi, M., Farazi, M., & Bakhshi, E. (2016). Evaluation of Effects of Gradual Increase Length and Complexity of Utterance (GILCU) treatment method on the reduction of dysfluency in school-aged children with stuttering. *Iranian Rehabilitation Journal*, *14*(1), 59–62. https://doi.org/10.15412/J.IRJ.08140109

Bauer, J. J. (2004). *Task dependent modulation of voice F0 responses elicited by perturbations in pitch of auditory feedback during English speech and sustained vowels* (Doctoral dissertation, Northwestern University).

Bauer, J. J., Hubbard Seery, C., LaBonte, R., & Ruhnke, L. (2007). Voice F0 responses elicited by perturbations in pitch of auditory feedback in individuals that stutter and controls. *The Journal of the Acoustical Society of America*, *121*(5), 3201–3201. https://doi.org/10.1121/1.4782465

Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *The Journal of the Acoustical Society of America*, *114*(2), 1048-1054. https://doi.org/10.1121/1.1592161

Bauer, J. J., Mittal, J., Larson, C. R., & Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude. *The Journal of the Acoustical Society of America*, *119*(4), 2363-2371. https://doi.org/10.1121/1.2173513

Black, J. W. (1951). The effect of delayed side-tone upon vocal rate and intensity. *Journal of Speech and Hearing Disorders*, *16*(1), 56-60. https://doi.org/10.1044/jshd.1601.56

Bosshardt, H. G., Sappok, C., Knipschild, M., & Hölscher, C. (1997). Spontaneous Imitation of Fundamental Frequency and Speech Rate by Nonstutterers and Stutterers. *Journal of Psycholinguistic Research*, *26*(4), 425–448. https://doi.org/10.1023/A:1025030120016

Bothe, A. K., Davidow, J. H., Bramlett, R. E., & Ingham, R. J. (2006). Stuttering treatment research 1970–2005: I. Systematic review incorporating trial quality assessment of behavioral, cognitive, and related approaches. *American Journal of Speech-Language Pathology*. https://doi.org/10.1044/1058-0360(2006/031)

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161. https://doi.org/10.1121/1.423073

Burnett, T. A., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, *112*(3), 1058–1063. https://doi.org/10.1121/1.1487844

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and Language*, *129*(1), 24–29. https://doi.org/10.1016/j.bandl.2014.01.002

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE*, *7*(7), 1–13. https://doi.org/10.1371/journal.pone.0041830

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *Journal of Neuroscience*, *31*(45), 16483–16490. https://doi.org/10.1523/JNEUROSCI.3653-11.2011

Cheadle, O., Sorger, C., & Howell, P. (2018). Identification of neural structures involved in stuttering using vibrotactile feedback. *Brain and Language*, *180–182*(May), 50–61. https://doi.org/10.1016/j.bandl.2018.03.002

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F0 responses to pitch-shifted voice feedback during English speech. *The Journal of the Acoustical Society of America*, *121*(2), 1157–1163. https://doi.org/10.1121/1.2404624

Costello-Ingham, J. C. (1993). Current status of stuttering and behavior modification—I: Recent trends in the application of behavior modification in children and adults. *Journal of Fluency Disorders*, *18*(1), 27-55.

Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S. E. (2018). Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. *Developmental Science*, *21*(2), 1–11. https://doi.org/10.1111/desc.12521

Donath, T. M., Natke, U., & Kalveram, K. T. (2002). Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *The Journal of the Acoustical Society of America*, *111*(1), 357–366. https://doi.org/10.1121/1.1424870

Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *Journal of the Acoustical Society of America*, *70*(1), 45–50. https://doi.org/10.1121/1.386580

Fairbanks, G. (1955). Selective vocal effects of delayed auditory feedback. *The Journal of Speech and Hearing Disorders*, *20*(4), 333–346. https://doi.org/10.1044/jshd.2004.333

Fucci, D., Petrosino, L., Gorman, P., & Harris, D. (1985). Vibrotactile magnitude production scaling: A method for studying sensory-perceptual responses of stutterers and fluent speakers. *Journal of Fluency Disorders*, *10*(1), 69–75. https://doi.org/10.1016/0094-730X(85)90007-5

Fucci, D., Petrosino, L., Schuster, S., & Belch, M. (1991). Lingual vibrotactile threshold shift differences between stutterers and normal speakers during magnitude-estimation scaling. *Perceptual and motor skills*, *73*(1), 55-62. https://doi.org/10.2466%2Fpms.1991.73.1.55

Gallop, R. F., & Runyan, C. M. (2012). Long-term effectiveness of the SpeechEasy fluency-enhancement device. *Journal of Fluency Disorders*, *37*(4), 334–343. https://doi.org/10.1016/j.jfludis.2012.07.001

Geetha, Y. V., Sangeetha, M., Sundararaju, H., Sahana, V., Akshatha, V., & Antonye, L. (2017). Effects of Altered Auditory and Oro-sensory Feedback on Speech Naturalness in Persons With and Without Stuttering. *Board of Reviewers*, 12.

Goldiamond, I. Stuttering and fluency as manipulatable operant response classes. In: Krasner, L.; Ullman, L., editors. Research in behavior modificaton. Holt, Rhinehart and Winston; 1965. p. 106-156.

Griffin, J. W. (1990). Basic pathologic processes in the nervous system. *Toxicologic Pathology*, *18*(1 II), 83–88. https://doi.org/10.1177/019262339001800113

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*(2), 133–141. https://doi.org/10.1007/s002219900237

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech and Hearing Research*, *37*(6), 1313–1319. https://doi.org/10.1044/jshr.3706.1313

Heinks-Maldonado, T. H., & Houde, J. F. (2005). Compensatory responses to brief perturbations of speech amplitude. *Acoustic Research Letters Online*, *6*(3), 131–137. https://doi.org/10.1121/1.1931747

Hendelman, W. (2015). *Atlas of functional neuroanatomy*. CRC press.

Hochman, G., & Erev, I. (2013). The partial-reinforcement extinction effect and the contingent-sampling hypothesis. *Psychonomic Bulletin and Review*, *20*(6), 1336–1342. https://doi.org/10.3758/s13423-013-0432-1

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216. https://doi.org/10.1126/science.279.5354.1213

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research*, *45*(2), 295–310. https://doi.org/10.1044/1092-4388(2002/023)

Howell, P. (2004a). Assessment of Some Contemporary Theories of Stuttering That Apply to Spontaneous Speech. *Contemporary Issues in Communication Science and Disorders*, *31*(Spring), 123–140. https://doi.org/10.1044/cicsd_31_s_123

Howell, P. (2004b). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary issues in communication science and disorders*, *31*(Spring), 123-140. http://www.speech.psychol.ucl.ac.uk/PAPERS/PDF/Howell1.pdf

Howell, P. (1990). Changes in voice level caused by several forms of altered feedback in fluent speakers and stutterers. *Language and Speech*, *33*(4), 325-338. https://doi.org/10.1177%2F002383099003300402

Howell, P. (2011). Listen to the lessons of The King's Speech: a film that shows King George VI struggling with a stammer could raise awareness and change treatments. *Nature*, *470*(7332), 7-8. https://doi.org/10.1038/470007a

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics*, *36*(3), 296–302. https://doi.org/10.3758/BF03206371

Howell, P., El-Yaniv, N., & Powell, D. J. (1987). Factors affecting fluency in stutterers when speaking under altered auditory feedback. In *Speech motor dynamics in stuttering* (pp. 361-369). Springer, Vienna.

Howell, P., & Powell, D. J. (1987). Delayed auditory feedback with delayed sounds varying in duration. *Perception & Psychophysics*, *42*(2), 166–172. https://doi.org/10.3758/BF03210505

Howell, P., & Sackin, S. (2000). Speech rate manipulation and its effects on fluency reversal in children who stutter. *Journal of Developmental and Physical Disabilities*, *12*, 291-315.

Howell, P., & Sackin, S. (2002). Timing interference to speech in altered listening conditions. *The Journal of the Acoustical Society of America*, *111*(6), 2842-2852.

Howell, P., Sackin, S., & Glenn, K. (1997a). Development of a two-stage procedure for the automatic recognition of dysfluencies in the speech of children who stutter: I. Psychometric procedures appropriate for selection of training material for lexical dysfluency classifiers. *Journal of Speech, Language and Hearing Research*, *40*, 1073-1084

Howell, P., Sackin, S., & Glenn, K. (1997b). Development of a two-stage procedure for the automatic recognition of dysfluencies in the speech of children who stutter: II. ANN recognition of repetitions and prolongations with supplied word segment markers *Journal of Speech, Language and Hearing Research*, *40*, 1085-1096.

Howell, P., & Williams, M. (1992). Acoustic analysis and perception of vowels in children's and teenagers' stuttered speech. *The Journal of the Acoustical society of America*, *91*(3), 1697-1706. https://doi.org/10.1121/1.402449

Howell, P., Wingfield, T., & Johnson, M. (1988). Characteristics of the speech of stutterers during normal and altered auditory feedback. Institute of Acoustics.

Howell, P., Young, K., & Sackin, S. (1992). Acoustical changes to speech in noisy and echoey environments. In *Speech Processing in Adverse Conditions*. 223-226. Retrieved from: https://www.isca-speech.org/archive_open/spac/spac_223.html

Iggo, A. (1986). Sensory receptors in the skin of mammals and their sensory function. *Pain*, *26*(1), 121. https://doi.org/ 10.1016/0304-3959(86)90184-3

Jones, J. A., & Munhall, K. G. (2002). The role of auditory feeback during phonation: Studies of Mandarin tone production. *Journal of Phonetics*, *30*(3), 303–320. https://doi.org/10.1006/jpho.2001.0160

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America*, *108*(3), 1246. https://doi.org/10.1121/1.1288414

Juliano, S. L., & Mclaughlin, D. F. (1999). Somatic senses 2: discriminative touch. *Neuroscience for rehabilitation (pp. 111-130), 2nd ed. Washington DC: Lippincott Williams & Wilkins*, 98-9.

Kalinowski, J., Armson, J., Stuart, A., & Gracco, V. L. (1993). Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech*, *36*(1), 1-16. https://doi.org/10.1177%2F002383099303600101

Kalinowski, J., Stuart, A., Sark, S., & Armson, J. (1996). Stuttering amelioration at various auditory feedback delays and speech rates. *International Journal of Language and Communication Disorders*, *31*(3), 259–269. https://doi.org/10.3109/13682829609033157

Kalveram, K. T., & Jäncke, L. (1989). Vowel duration and voice onset time for stressed and nonstressed syllables in stutterers under delayed auditory feedback condition. *Folia Phoniatrica*, *41*(1), 30-42.

Kuniszyk-Józkowiak, W., & Adamczyk, B. (1989). Effect of tactile echo and tactile reverberation on the speech fluency of stutterers. *International Journal of Rehabilitation Research*, *12*(3), 312-317. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=ovfta&NEWS=N&AN=00004356-198909000-00009.

Kuniszyk-Jóźkowiak, W., Smołka, E., & Adamczyk, B. (1996). Effect of acoustical, visual and tactile echo on speech fluency of stutterers. *Folia phoniatrica et logopaedica*, *48*(4), 193-200. https://doi.org/10.1159/000266408

Kuniszyk-Jóźkowiak, W., Smołka, E., & Adamczyk, B. (1997). Effect of acoustical, visual and tactile reverberation on speech fluency of stutterers. *Folia Phoniatrica et Logopaedica*, *49*(1), 26–34. https://doi.org/10.1159/000266434

Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of neuroscience*, *32*(27), 9351–9358. https://doi.org/10.1523/JNEUROSCI.0404-12.2012

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, *14*(4), 677-709. https://doi.org/10.1044/jshr.1404.677

Larson, C. R., Sun, J., & Hain, T. C. (2007). Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control. *The Journal of the Acoustical Society of America*, *121*(5), 2862–2872. https://doi.org/10.1121/1.2715657

Lee, B. S. (1950). Effects of Delayed Speech Feedback. *Journal of the Acoustical Society of America*, *22*(6), 824–826. https://doi.org/10.1121/1.1906696

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders*, *31*(2), 71–89. https://doi.org/10.1016/j.jfludis.2006.04.001

Liu, H., & Larson, C. R. (2007). Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *The Journal of the Acoustical Society of America*, *122*(6), 3671–3677. https://doi.org/10.1121/1.2800254

Lotzmann, G. (1961). On the use of varied delay times in stammerers. *Folia phoniatrica*, *13*, 276-310.

Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language and Communication Disorders*, *47*(4), 451–456. https://doi.org/10.1111/j.1460-6984.2011.00111.x

Macleod, J., Kalinowski, J., Stuart, A., & Armson, J. (1995). Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. *Journal of Communication Disorders*, *28*(3), 217–228. https://doi.org/10.1016/0021-9924(94)00010-W

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or Insufficiently Activated Internal Models and Feedback-Biased Motor Control as Sources of Dysfluency: A Theoretical Model of Stuttering. *Contemporary Issues in*

*Communication Science and Disorders*, *31*(Spring), 105–122. https://doi.org/10.1044/cicsd_31_s_105

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *The Journal of the Acoustical Society of America*, *113*(3), 1587–1593. https://doi.org/10.1121/1.1543928

Natke, U., Grosser, J., & Kalveram, K. T. (2001). Fluency, fundamental frequency, and speech rate under frequency-shifted auditory feedback in stuttering and nonstuttering persons. *Journal of Fluency Disorders*, *26*(3), 227-241. https://doi.org/10.1016/S0094-730X(01)00099-7

Natke, U., & Kalveram, K. T. (2001). Effects of Frequency-Shifted Auditory Feedback on Fundamental Frequency of Long Stressed and Unstressed Syllables. *Journal of Speech, Language, and Hearing Research*, *44*(3), 577–584. https://doi.org/10.1044/1092-4388(2001/045)

Nessel, E. (1958). Die verzögerte Sprachrückkopplung (Lee-Effekt) bei Stotterern. *Folia Phoniatrica et Logopaedica*, *10*(4), 199-204.

Novak A. The influence of delayed auditory feedback in stutterers. *Folia Phoniatrica*. 1978; *30*, 278– 285.

O'Donnell, J. J., Armson, J., & Kiefte, M. (2008). The effectiveness of SpeechEasy during situations of daily living. *Journal of Fluency Disorders*, *33*(2), 99–119. https://doi.org/10.1016/j.jfludis.2008.02.001

Pollard, R., Ellis, J. B., Finan, D., & Ramig, P. R. (2009). Effects of the speechEasy on objective and perceived aspects of stuttering: A 6-month, phase i clinical trial in naturalistic environments. *Journal of Speech, Language, and Hearing Research*, *52*(2), 516–533. https://doi.org/10.1044/1092-4388(2008/07-0204)

Purcell, D. W., & Munhall, K. G. (2006a). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America*, *120*(2), 966–977. https://doi.org/10.1121/1.2217714

Purcell, D. W., & Munhall, K. G. (2006b). Compensation following real-time manipulation of formants in isolated vowels. *The Journal of the Acoustical Society of America*, *119*(4), 2288–2297. https://doi.org/10.1121/1.2173514

Reed, P., & Howell, P. (2000). Suggestions for Improving the Long-term Effects of Treatments for Stuttering: A Review and Synthesis of Frequency-shifted Feedback and Operant Techniques. *European Journal of Behavior Analysis*, *1*(2), 89–106. https://doi.org/10.1080/15021149.2000.11434158

Ritto, A. P., Juste, F. S., Stuart, A., Kalinowski, J., & de Andrade, C. R. F. (2016). Randomized clinical trial: the use of SpeechEasy® in stuttering treatment. *International Journal of Language and Communication Disorders*, *51*(6), 769–774. https://doi.org/10.1111/1460-6984.12237

Rosen, S., & Howell, P. (2011). *Signals and systems for speech and hearing* (Vol. 29). Brill.

Ryan, B. (1974). *Programmed therapy for stuttering in children and adults*. Thomas.

Sares, A. G., Deroche, M. L. D., Shiller, D. M., & Gracco, V. L. (2018). Timing variability of sensorimotor integration during vocalization in individuals who stutter. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-34517-1

Siegel, A., & Sapru, H. N. (2006). *Essential neuroscience*. Lippincott Williams & Wilkins.

Sivasankar, M., Bauer, J. J., Babu, T., & Larson, C. R. (2005). Voice responses to changes in pitch of voice or tone auditory feedback. *The Journal of the Acoustical Society of America*, *117*(2), 850-857. https://doi.org/10.1121/1.1849933

Soderberg, G. A. (1969). Delayed auditory feedback and the speech of stutterers: A review of studies. *Journal of Speech and Hearing Disorders*, *34*(1), 20-29.

Sparks, G., Grant, D. E., Millay, K., Walker-Batson, D., & Hynan, L. S. (2002). The effect of fast speech rate on stuttering frequency during delayed auditory feedback. *Journal of Fluency Disorders*, *27*(3), 187–201. https://doi.org/10.1016/S0094-730X(02)00128-6

Stuart, A., & Kalinowski, J. (2004). The perception of speech naturalness of post-therapeutic and altered auditory feedback speech of adults with mild and severe stuttering. *Folia Phoniatrica et Logopaedica*, *56*(6), 347–357. https://doi.org/10.1159/000081082

Stuart, A., Kalinowski, J., Armson, J., Stenstrom, R., & Jones, K. (1996). Fluency effect of frequency alterations of plus/minus one-Half and one-quarter octave shifts in auditory feedback of people who stutter. *Journal of Speech, Language, and Hearing Research*, *39*(2), 396–401. https://doi.org/10.1044/jshr.3902.396

Stuart, A., Kalinowski, J., Rastatter, M. P., Saltuklaroglu, T., & Dayalu, V. (2004). Investigations of the impact of altered auditory feedback in-the-ear devices on the speech of people who stutter: Initial fitting and 4-month follow-up. *International Journal of Language and Communication Disorders*, *39*(1), 93–113. https://doi.org/10.1080/13682820310001616976

Stuart, A., Kalinowski, J., Saltuklaroglu, T., & Guntupalli, V. (2006). Investigations of the impact of altered auditory feedback in-the-ear devices on the speech of people who stutter: One-year follow-up. *Disability and Rehabilitation*, *28*(12), 757–765. https://doi.org/10.1080/09638280500386635

Stuart, A., Xia, S., Jiang, Y., Jiang, T., Kalinowski, J., & Rastatter, M. P. (2003). Self-contained in-the-ear device to deliver altered auditory feedback: Applications for stuttering. *Annals of Biomedical Engineering*, *31*(2), 233–237. https://doi.org/10.1114/1.1541014

Snyder, G. J., Waddell, D., Blanchet, P., & Ivy, L. J. (2009). Effects of digital vibrotactile speech feedback on overt stuttering frequency. *Perceptual and motor skills*, *108*(1), 271-280.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and cognitive processes*, *26*(7), 952-981. https://dx.doi.org/10.1080%2F01690960903498424

Van Borsel, J., Reunes, G., & Van Den Bergh, N. (2003). Delayed auditory feedback in the treatment of stuttering: Clients as consumers. *International Journal of Language and Communication Disorders*, *38*(2), 119–129. https://doi.org/10.1080/1368282021000042902

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319. https://doi.org/10.1121/1.2773966

Waddell, D. E., Goggans, P. M., & Snyder, G. J. (2012). Novel tactile feedback to reduce overt stuttering. *NeuroReport*, *23*(12), 727–730. https://doi.org/10.1097/WNR.0b013e328356b108

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain : A Journal of Neurology*, *131*(Pt 1), 50–59. https://doi.org/10.1093/brain/awm241

Weidig, T. (2005). The statistical fluctuation in the natural recovery rate between control and treatment group dilutes their results. Rapid response to "Randomised controlled trial of the Lidcombe programme of early stuttering intervention" by Jones, Onslow, Packman, Williams, Ormond, Schwarz & Gebski (2005). *British Medical Journal* [Announcement posted on the World Wide Web]. 27-11-2009, from: http://bmj.bmjjournals.com/cgi/eletters/331/7518/659#115238