

Published in final edited form as:

*Stammering Res.* 2004 April 1; 1(1): 31–46.

## Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering

**Peter Howell**

Department of Psychology Centre for Human Communications Institute of Cognitive Neuroscience And Institute of Movement Neuroscience University College London Gower St., London WC1E 6BT

### Abstract

It has been known for at least a hundred years that the speech of a person who stammers becomes more fluent when alterations are made to the speaking environment. Alterations that lead to an improvement in fluency include a) noises that prevent a speaker hearing his or her own voice, and b) manipulations to the sound of a speaker's voice before it is heard. Examples of manipulations that have been made are introducing a delay, and shifting the voice up or down in frequency. The influences all these alterations have on fluent speakers and speakers who stammer, that have been established over the last century, are reviewed. In addition, the ways in which these phenomena have been explained for both fluent speaker and speakers who stammer are outlined. Several previous findings have potential significance for ways in which the fluency-enhancing effects of these alterations in speakers who stammer could be employed in clinical settings. These are highlighted and discussed, mainly in connection with the SpeechEasy™ prosthetic device for treating stammering.

### Keywords

Altered auditory feedback; delayed auditory feedback; frequency shifted feedback; EXPLAN; mirror neurons; SpeechEasy™

### Motivation for this review

Interest in the effects of altering the speaking environment of speakers who stammer is currently at an all-time high. This is largely due to the publicity that the SpeechEasy™ in the ear stammering aid has received: The fluency-enhancing effects of this aid have been demonstrated to dramatic effect on the Oprah Winfrey show, and the device has featured on the front page of USA today. SpeechEasy™ alters the sound of the speaker's voice before he or she hears it in one of two ways: 1) by delaying it, or 2) by shifting the speech spectrum (frequency shifting). The former creates a speaking situation like that in an echoey auditorium, and the latter gives the speaker the impression of speaking at the same time as another speaker (either one with a deeper voice or one with a higher voice, depending on which way the speech spectrum is shifted). Examination of these effects in speakers who stammer was initiated by Lee (1951) for delaying, and by Howell, El-Yaniv and Powell (1987) for frequency shifting.

The favorable storm of publicity has met with a more cautious response by some professionals involved in delivering treatment. For instance, the paucity of research about the device led Roger and Janis Ingham to point out in a recent letter to the American Speech Language and Hearing Association's leader magazine that there is no evidence-based practice that SpeechEasy™ “produces any sustained and satisfactory improvements in fluency”. In a response to this letter, Greg Snyder raised the issue of whether it is appropriate to delay introduction of the device until such time-consuming and costly research has been conducted. The rights and wrongs of each of these positions is not one that will be quickly resolved so, though it has been aired here, it will not feature directly in the remainder of this review. As currently there are such strongly held positions about fluency-enhancing aids, the time seems right to review their history, comment on their pros and cons, see how they might be integrated with other forms of treatment and speculate about the ways in which use of such aids may advance in the near future.

## Definitions

As described above, the SpeechEasy™ equipment is a portable device that implements procedures known to improve the fluency of people who stammer. Delaying and frequency shifting are two techniques often referred to generically as altered auditory feedback procedures. Auditory feedback is a value-laden term that carries the implicit idea that speakers listen to the sound of their voice and send the result of this processing back through the brain to a level where this information can be compared with the production the speaker intended to produce. If the sound heard was the one intended, then speech was fluent. If the intended sound was different to what the speaker heard, an error has crept into the process of speech production. Corrective action can then be taken. This whole process is one of negative, or compensatory, feedback. The overall process (using feedback to determine whether an error has occurred, and then acting on it) is referred to as monitoring. Though it is conceivable that the process of speech control works like this, other explanations are possible. To admit these possibilities, a more neutral term is needed. Hence, ‘alterations to recurrent auditory information’ (ARAI) is used in preference to ‘altered auditory feedback’. ARAI covers both feedback and non-feedback interpretations of the effects that occur when the auditory environment is altered. This term will be used when referring to the several methods of making alterations. The terms delayed auditory feedback (DAF) and frequency shifted feedback (FSF) also beg the question of whether the effects are a result of feedback or not. However these terms will be employed in this review because they are so widely used in the literature.

## Structure of the review

There is no doubt that if the listening conditions change in the ways mentioned above while a person who stammers is speaking, their speech control improves. Investigation into the effects of such ARAI can be divided roughly into four historical stages, characterized in terms of what equipment was available. The stages are: 1) before any equipment was effectively available; 2) electrical hardware; 3) cheap programmable computers; and 4) portable microelectronic devices. The overriding questions at each historical stage are: 1) whether the advantageous effects of artificially manipulating what speakers who stammer hear can be employed in treatment (practice); and 2) what this indicates about the nature of stammering (theory). While the discussion in the first three stages seems fairly uncontroversial, the theory section in stage four selects two theories developed to account, *inter alia*, for why FSF improves the speech of speakers who stammer. One of these theories is EXPLAN that was developed by the author of this article. The other theory (authored by a group at East Carolina University) offers a contrasting account of some of the same effects that are addressed by EXPLAN. “Stammering Research” is intended to promote discussion

on practical and theoretical topics about stammering and allied issues. Thus in this part of stage four, I argue against the East Carolina theory and present evidence in favor of the 'home theory' (EXPLAN). Undoubtedly, the Carolina group, as well as other interested parties, will wish to address alternative positions through the open peer commentary format of the journal. The article finishes off with some speculation about future prospects concerned with ARAI.

## Stage 1

### Empirical observations on the effects of speaking in noise by people who stammer

Work on ARAI started with the observations, made by people who stammer, that speaking in noisy environments improved their voice control. This result must have been startling as there was no literature that would allow them to understand how a speech production problem could be affected by what you hear. These effects were only experienced adventitiously by isolated individuals as there was no equipment available that allowed the effects to be manipulated and investigated in a controlled way. In the first published experimental study that I have been able to locate, Kern (1932) used a Barany drum as a noise source to study this phenomenon.

One issue that was addressed as a result of these early observations was that if stammering is a result of a hearing deficit, the problem should stop if hearing is lost. Contrary to this prediction studies at this time showed loss of hearing to be associated with onset (not cessation) of stammering in some individuals (Albright & Malone, 1942; Backus, 1939).

### Empirical observations on the effects of speaking in noise in fluent speakers

One other topic that predated experimental work on ARAI that is relevant for later, concerns the influence of speaking when the voice is amplified (Fletcher, Raff & Parmley, 1918) or when noise is present (Lombard, 1911). Speakers who stammer change their voice level in the same direction as fluent speakers when noise is present and when their voice is amplified or attenuated (Howell, 1990). When voice level is amplified, speakers reduce their voice level and when voice level is reduced, speakers increase their voice level (called the Fletcher effect). Conversely, when noise level increases, speakers increase their voice level and when noise level reduces, speakers reduce their voice level (called the Lombard effect). It is possible that these compensations could be the result of a negative feedback mechanism for regulating voice level. If speakers need to hear their voice to control it but cannot do so, either because noise level is high or voice level is low, they compensate by increasing level. Speakers would compensate in the opposite way if their speech is too loud (low noise level or when the voice is amplified). Note, however, that explanations other than a feedback account, are also possible (see, for instance, Lane and Tranel, 1971 who discuss the view that voice level changes are made so that the audience, rather than the speaker himself or herself, does not receive speech at too high or too low a level).

## Stage 2

### Empirical observations on the effects of speaking under echo in fluent speakers

In the 1950s, the rapid growth in telephone use caused engineers to become interested in how alterations affected fluent speakers' speech control. Telephones can transmit a limited range of frequencies, the equipment can introduce delays and the voice can be masked by noise, and voice level changes can occur. Thus telephones create ARAI and telephone companies needed to know how speech was affected. Most attention at this time and subsequently has been on the effects of delay (CCITT, 1989a, 1989b) and is an on-going problem since the introduction of cellular phones and satellite technology. Speaking along

with a delayed version of the voice (DAF) caused drawing (usually on the medial vowels), led to a Lombard effect (increased voice level), while pitch became monotone, speech errors arose and messages took longer to complete than messages produced in normal listening conditions (Fairbanks, 1955).

### **Theoretical accounts of the effects of speaking under echo in fluent speakers**

These observations led to various versions of ‘feedback’ theory (Black, 1951; Lee, 1950). The essential feature of these theories is that the current speech output is sent back to a sensing device that controls future output (Brown & Campbell, 1948). The information that arises at this sensing device is used to correct an activity when it exceeds predetermined limits. In the case of DAF procedures, the sound of a speaker’s voice is transformed by delaying before it reaches the sensing device, so the segment of speech that is heard at a particular time is different from the segment that the speaker intended to produce at that time. A feedback monitoring explanation maintains that this discrepancy is detected and the corrections the speaker then makes, introduce, rather than remove, errors. If this interpretation is correct, then the delays at which errors are observed, indicate what segments are involved in speech control. The notion behind this is that a delay equal to the length of the unit used for output, results in the speaker getting feedback about the preceding segment when he or she is producing the next segment. Using this idea, Black (1951) argued that since a delay of 200 ms is most disruptive on speech control and that as this corresponds roughly with the length of a syllable, then the unit used by speakers to monitor feedback is the syllable.

### **Empirical observations on the effects of speaking under echo in speakers who stammer**

When DAF was presented to people who stammer, fluency was found to improve (as had been reported earlier when a noise masked these speakers’ speech). Researchers who investigated the fluency-enhancing effects of DAF on people who stammer in the 1950s and 1960s include Nessel (1958), Soderberg (1960), Chase, Sutton and Rapin (1961), Lotzmann (1961), Neelley (1961), Goldiamond (1965), Ham and Steer (1967) and Curlee and Perkins (1969). Stimulated by the findings of these early investigators, several portable maskers and DAF devices were developed.

Following the pioneering work of Goldiamond (1965), DAF was introduced into an influential treatment program by Ryan (1974). DAF was initially presented with a delay long enough to produce slowing of speech (based on the work on fluent speakers mentioned above, most slowing would occur when speech is delayed by 200 ms). The delay was faded over a series of test sessions so that rate was reestablished to normal limits, hopefully with some retention of the fluent patterns established when speech rate was slow. As recently as 1993, Costello-Ingham also maintained that the only function of DAF was to control speech rate. As she put it: “The functional variable in regard to the reduction of stuttering is not DAF, but prolonged speech, and the latter can be produced without reliance on a DAF machine” (Costello-Ingham, 1993, p.30).

Other techniques for treating stammering, not involving ARAI, were investigated at this time. One that deserves special mention is the Lidcombe learning procedure, because of its current popularity and some comments are made under “future possibilities” about how DAF or FSF could feature in a modification of such an operant procedure. Onslow, Andrews and Lincoln (1994) describe the technique as follows. It “is an operant treatment that incorporates parental verbal contingencies for stuttered speech and stutter-free speech. The contingencies for stutter-free speech are praise and tangible reinforcement, and the contingencies for stuttering are that the parents identify a stuttered utterance and request the child to correct the utterance.”

A further important claim that was made at this time that was embraced by several eminent workers was that DAF produces similar effects in fluent speakers to those that people who stammer ordinarily experience – in particular drawling and speech errors. This prompted Lee (1951) to refer to DAF as a form of “simulated” stammer. In an extension of this point of view, Cherry and Sayers (1956) used DAF as a way of simulating stammering in fluent speakers to establish the basis of the problem. They extracted two different sources of sound that are heard whilst speaking normally (the sound transmitted over air and that transmitted through bone). They then examined which of these ‘feedback’ components led to increased stammering rates in fluent speakers when each of them was delayed. The bone-conducted component seemed to be particularly effective in increasing ‘simulated’ stammering and they proposed that this source of feedback also led to the problem in speakers who stammer. They then designed a therapy that involved playing noise to speakers who stammer that was intended to mask out the problematic bone-conducted component of vocal ‘feedback’. They reported that fluency improved when the voice was masked in this way.

In another particularly imaginative study, Sutton and Chase (1961) manipulated when noise was on or off using a voice-activated relay while subjects read aloud. They compared the fluency-enhancing effects of noise that was on continuously, noise that was presented only while the speaker was speaking and noise presented only during the silent periods between speech. They found all these conditions were equally effective. It appears from this that the operative effect is not simply masking as there is no sound to mask when noise is presented during silent periods. However, Webster and Lubker (1968a) pointed out that voice-activated relays take time to operate and so some noise would have been present at the onset of words. Therefore a masking effect cannot be ruled out.

### **Theoretical accounts that suggested a sensory deficit in people who stammer**

Theorists at this time were proposing that malfunction in different parts of the auditory system might offer an account of stammering. Webster and Lubker, (1968b) for instance, postulated that middle ear muscle contraction in speakers who stammer disrupts the auditory feedback that they receive. Whenever the middle ear muscles contract, the middle ear system increases impedance to sound transmission. The muscles contract prior to vocalization, resulting in attenuation and low-pass filtering of the speech (Teig, 1973). Shearer (1966) reported that the timing of this muscle activity is abnormal in speakers who stammer. According to Webster and Lubker's theory, the abnormal contraction and relaxation of the middle ear muscles of the person who stammers would produce abnormal speech feedback of fluctuating intensity that leads to speech control problems. The positive effects of DAF on speakers who stammer could then arise because this form of ARAI keeps the muscles constantly contracted and removes the fluctuating auditory feedback that created the problem.

## **Stage 3**

### **Conceptual and empirical problems for a feedback account of fluent speech control**

Though in the previous period Lee, and Cherry and Sayers were interested in speech control of fluent speakers and speakers who stammer, the 1970s and 1980s started to see some division between people interested in fluent speech control and those interested in stammering. Generally speaking, a ‘feedback’ process as candidate for explaining speech control process was dropped in fluent speech, but was retained by people interested in how people who stammer control their voice. Thus, work on fluent speech, including papers by Borden (1979), Howell, Powell and Khan (1983), and Lane and Tranel (1971), began to question feedback interpretations of the effects of ARAI, and alternative accounts were

proposed. There were both conceptual and empirical objections that led to rejection of the view that ARAI is used as sensory feedback to linguistic planning mechanisms.

Borden (1979) discussed several conceptual issues for a feedback point of view. One question she raised was how quickly information can be recovered from the auditory signal. Auditory processing time is estimated to take around 100-200 ms. Auditory output from any segment around this duration would reach the feedback mechanism too late to be used for control of its own segment. A second question she raised was based on the observation that speakers with hearing impairment, who had established language before they sustained their loss, can continue to speak. This suggests that speech can proceed without sensory feedback.

A further conceptual problem is that the amount of phonetic information a speaker can recover about vocal output is limited because bone-conducted sound masks a speaker's phonetic output (see Howell and Powell, 1984 for a study on this issue and Howell, 2002, for an extended discussion of the problems this raises for feedback accounts). Degradation of the sound of the voice would limit the usefulness of the feedback that a speaker can recover by listening to his or her own voice, making it an unlikely source of information for use for feedback control.

One question that arises if the sound of the voice does not contain phonetic information, is whether the delayed sound during DAF has to be speech to produce the disruptions to fluent speakers' speech? Howell and Archer (1984) addressed this question by transforming speech into a noise that had the same temporal structure as speech, but none of the phonetic content. Then they delayed the noise sound and compared performance of this with performance under standard DAF. The two conditions produced equivalent disruption over a range of delays. This suggests that the DAF signal does not need to be a speech sound to affect control in the same way as observed under DAF, and indicates that speech does not go through the speech comprehension system before it can be used as feedback. The disruption could arise, however, if asynchronous inputs affect operation of lower level mechanisms involved in motor control.

### **Revisions in theory in response to the problems for a feedback account of fluent speech control**

The above arguments and Howell and Archer's (1984) experimental evidence, undermine the case for auditory feedback monitoring in fluent speakers. There have been several reactions: 1) Some have argued for an auditory feedback processing mechanism that operates at the prosodic level (Donath, Natke & Kalveram, 2002; Kalveram, 2001; Kalveram & Jaencke, 1989). Prosodic processes operate over long time periods. Thus, the problem of obtaining auditory feedback early enough would not be such a problem if prosodic units are used for feedback control as it is for the view that syllables are the unit that is used. 2) Borden (1979) argued that auditory feedback is used in circumscribed situations. These include when language is being acquired (either developmentally or as a second language in adulthood), and when the speaker's voice is altered. 3) Howell et al., (1983) developed a non-feedback account of the particular effects of DAF. Lane and Tranel (1971) offered a non-feedback account of the effects of alterations to voice level that were described earlier in this review. 4) Some authors adopted feedforward, instead of feedback, models (Kawato, Furakawa & Suzuki, 1987). These models maintain that movement errors are continuously computed and used (when they arise) as correction signals. They get round the problem of feedback being slow by doing the work in advance of the movement. Such a model has been applied to one of the situations Borden (1979) regarded as reliant on auditory feedback (developmental speech acquisition) by Guenther (2001).



Howell et al.'s (1983) account has particular relevance to the effects of ARAI on speakers who stammer because it involved DAF that improves the fluency of these speakers. It is worth giving a little of the background detail of this account (their disruptive rhythm hypothesis, DRH). The basic issue addressed by DRH was how to account for the disruptive effects of DAF if, as Howell and Archer's (1984) results indicate, ARAI does not send information through the speech perception system to provide information to reinitiate speech when it is in error. From a rhythmic perspective, DAF involves speaking one utterance while hearing another that is out of synchrony with it (in contrast with normal listening where the sound that is heard has a rhythm in synchrony with speech). Howell et al. (1983) considered two situations involving voice control to argue that synchronous activities are easy to perform and asynchronous ones are difficult. Canon singing is easy (as shown by the fact that it is one of the first forms of song that children are taught). There is also a form of medieval song, called hoquetus, that involves each singer producing a note synchronized to the offset of another singer's note. This form of singing is difficult to master. Canon singing points to the fact that it is easy to produce synchronous activities whether or not those activities contain any information about the speaker's own speech. The case of hoquetus shows that asynchronous activities (again, whether or not those activities contain any information about the speaker's own speech) are difficult and, by analogy, suggests that this is why DAF causes difficulties in speech control. In hoquetus, one singer's note finishes as the next speaker's note commences. This would correspond to the DAF situation in which speech is delayed by the length of the note, which would be the length of a syllable for notes a syllable in length. As observed earlier, a delay equal to the length of a syllable is maximally disruptive in DAF. DRH suggests that this delay is most disruptive because of the rhythmic relationship between what is heard and what is spoken, rather than because feedback about the wrong syllable is sent when this delay is used (as in traditional accounts).

### **Practical development of ARAI devices for speakers who stammer and some limitations about the fluency of the speech produced when using these devices**

Part of the growth in popularity of DAF as parts of treatment programs stemmed from the early claim by Lee (1951) (also endorsed by Cherry and Sayers, 1956), that DAF has the opposite effect on fluency between people who stammer and fluent speakers. This implies that DAF produces fluent speech in people who stammer. Considering first the effects of DAF on fluent speakers, the most notable effect is lengthening of medial vowels. Though these seem superficially similar to the prolongations people who stammer show, there are two differences that indicate this is more apparent than real: First, speakers who stammer have problems on consonants, not vowels (Howell, Wingfield & Johnson, 1988). Second, the consonants are in the initial position in an utterance (Wingate, 2002), not the medial position that the vowels occupy. The difference in distribution and phoneme type of the sounds that are elongated between DAF-speech in fluent speakers and prolongations that people who stammer produce, undermines the claim for complementarity between these two forms of speech.

A further point investigated at this time was whether people who stammer only lose disfluencies or whether they also show effects like fluent speakers. Howell et al. (1988) reported that people who stammer lose disfluencies under DAF but they also elongate the vowels (as do fluent speakers under DAF). These effects can be ameliorated by, for example, using short DAF delays (Kalinowski, Stuart, Sark & Armson, 1996), though standard equipment at this time usually limited the alterations that could be made to long delays. The difference between 'DAF-simulated', and true, stammering undermines the explanatory basis of Cherry and Sayers' (1956) work that led to masking therapy (though not the effectiveness of masking therapy itself). If Costello-Ingham's (1993) point of view that DAF is just a way of slowing speech that reduces stammering, and if DAF can be faded out

(as in Ryan, 1974) the side effects of DAF would not matter. However, other authors such as Novak (1978) have reported that the after-effects of DAF (vowel lengthening) persist into post treatment speech, so would affect speech communication adversely. One other objection about DAF is that it presents no sound at word onset, which is mostly the place where people who stammer have problems (Wingate, 2002). Lack of an altered sound at onset of syllables may explain why DAF has more effect on the medial vowels than initial consonants.

In the UK, development of two portable devices that included sensible design ideas was taking place. These were, 1) the Edinburgh masker pioneered by the stammering research unit at Edinburgh University (Dewar, Dewar, Austin & Brash, 1979) and, 2) the Hector aid designed and built by Ron Turrell and Graham Parkhouse with support from the forerunner of the British Stammering Association.

The Edinburgh masker consists of a microphone that is held on the larynx by a velcro band, a control box that is discretely hidden by the user (e.g. in the pocket) which is connected by plastic tubing to ear tips that the speaker inserts into the ear canal. The throat microphone detects voiced sounds, the control box triggers the masking noise (a low frequency buzz) that is delivered to the speaker's ears. The device has the advantage that the masking sound only occurs while the speaker is speaking, thus limiting the occasions on which the aid operates to the periods where the speaker may have problems. However, there are several drawbacks. First, the attachment of the microphone and the ear-inserts are somewhat unsightly and may be cosmetically unacceptable to wearers (particularly adolescents). Second, as the manufacturers of the device acknowledge in their instructions for users, the laryngeal microphone does not always trigger on initial parts of sounds, as for instance in words starting with low amplitude voiceless sounds. As most stuttering occurs on the initial sounds in an utterance (Wingate, 2002), the device does not always operate at the point at which speakers need assistance. As noted above, this was also a problem in Sutton and Chase's (1961) onset masker. The manufacturers of the Edinburgh masker suggest that speakers prelude speech attempts by saying 'm', 'er' or 'ah' that triggers the device to deliver a masking noise. However, the advisability of doing this is questionable as this strategy would substitute one unusual pattern of speech for another. This would be problematic in that work with DAF suggests that some of the odd patterns that arise with this ARAI persist into post-treatment speech (Novak, 1978) and the same could apply to speech produced under masking. Also, if the crucial factor that leads to DAF effects is delayed rhythm (Howell et al., 1983), then the Edinburgh masker with its inbuilt delay would work like DAF and produce speech with unwanted side effects. Third, again as the manufacturers acknowledge, the device produces a Lombard effect (a raising of the voice level). Once again this leads to unnatural sounding, in this case shouted, speech. Fourth, the insert earphones prevent speakers hearing outside sounds and this could potentially be dangerous if, for example, the masker is worn in the street (this is also a problem for the SpeechEasy™ device).

The Edinburgh masker was more popular, and its effects on fluency studied more extensively (Dewar et al., 1979), than the Hector aid. However, the Hector aid had some revolutionary characteristics behind its design that current ARAI technology ought to take on board (see future prospects for ideas on how this could be achieved). As far as I am aware, there has been no formal report describing the device or reporting on its effectiveness, apart from a single case study by Celia Levy who worked with a client over a period of eight weeks. This description relies mainly on that report and my own recollections of the device. The device consisted of a box with audio inputs and a vibrator output. The electronics measured speech rate using the audio input. The vibrator switched on if speech rate was outside acceptable speech rate ranges and signaled the speaker to slow



speech down. Presumably the imposed speech rate is the “bullying” which gave the aid its name ‘Hector’. Though rate control is not a form of ARAI, it is a form of feedback. Its primary attraction is that it targeted its indications that a speech rate change is needed on the episodes where stammering rate is likely to be highest, i.e. the fast rate sections (Howell, Au-Yeung & Pilgrim, 1999). This takes the idea of targeting feedback on sections that are problematic (Howell, El-Yaniv & Powell, 1987) a step further. Furthermore, if alterations are made intermittently (as in the Hector aid), they would cause less of a problem when worn in everyday speaking situations (see the above discussion about wearing the Edinburgh masker or SpeechEasy™ device in the street). Whether Hector works or not depends on the assumption that rate control is behind the problem that a person who stammers experiences (as Costello-Ingham, 1993, argued). As with the Edinburgh masker, the device has drawbacks. First, to be worn discretely, some adjustment to clothing was necessary (as noted in Levy's report of work with her patient). Second, when I made some measurements on the device in the 1980s, it did not track speech rate very accurately.

### **Empirical work rejecting theoretical accounts that suggested a sensory deficit in people who stammer**

As indicated, some workers proposed that stammering could arise as a result of an auditory (pure sensory) deficit at stage two. The two specific proposals made were that people who stammer have problems in dealing with bone-conducted sound (Cherry & Sayers, 1956) or that problems arise because the middle ear structures of speakers who stammer cannot transmit sound in the same way that those of fluent speakers do (Webster & Lubker, 1968b).

Cherry and Sayers' argument for problems in the bone-conducted route was based on the assumed similarity of stammered speech to DAF-speech in fluent speakers. Empirical studies that show that this is not so were reviewed above. Therefore, there is no basis to conclude that because sound delayed and transmitted through bone is more disruptive to fluent speakers than sound delayed and transmitted through air, speakers who stammer have problems dealing with sound transmitted through bone. Also, Howell and Powell (1984) compared Cherry and Sayers (1956) bone-conducted sound with actual bone-conducted sound and found marked differences. Cherry and Sayers' experimental manipulation created a sound that, though successful at disrupting fluent speech control, was nothing like bone-conducted sound. Once again this result shows that there are no grounds for concluding that speakers who stammer have problems in dealing with sound transmitted through bone.

The proposal that speakers who stammer have problems in transmitting sound through the middle ear system also failed empirical tests. Shearer's (1966) original work included very limited amounts of data. In an extensive study, Howell, Marchbanks and El-Yaniv (1986) were unable to find differences in middle ear operation between people who stammer and fluent controls (both during listening tests and during vocalization). Abnormal middle ear muscle operation seems, then, an unlikely basis for explaining the disorder.

## **Stage 4**

### **Empirical work when technology allowed an increased range of ARAI**

The advent of cheap computer power opened up possibilities for extending the type of alterations that can be made. The SpeechEasy™ device drew on the results of this work in terms of the alterations that it includes (DAF and FSF that improve fluency) and the operating ranges (delays and frequency shifts it is possible to make). These and other alterations that were explored are summarized next.

Howell and co-workers began to examine the implications of DRH for the effects of new forms of ARAI in people who stammer. They investigated the effects of various forms of

synchronous and asynchronous rhythms on the speech of people who stammer. One investigation on synchronous rhythms by Howell and El-Yaniv (1987), examined a metronome click that was automatically triggered by speech so that it was located at the onset of each syllable in the spontaneous speech of speakers who stammer. They found such a speech-synchronous metronome click was as effective at increasing fluency as an externally paced metronome. This suggests the effect of this novel metronome stimulus is not due to rate pacing (the speaker is free to adopt whatever rate he or she is comfortable with) and may be a result of having a click in synchrony with speech.

Howell et al. (1983) in the paper that introduced the DRH, pointed out that interrupting speech (by gating it on and off) produced asynchronous ARAI similar in some respects to what they considered to occur under DAF (disruption to rhythm, without any part of speech being delayed). They found some similarities between speech performance under interruption and DAF in fluent speakers. This manipulation remains to be investigated in people who stammer, but DRH predicts that it would lead to similar effects on fluency as DAF.

Howell, El-Yaniv and Powell (1987) created a frequency-shifted version of the speaker's voice that was synchronous with the speaker's voice. These authors used a speed-changing method (that produces a frequency shift in the same way that playing a tape recorder at different speeds does). To avoid the altered sound getting out of synchrony with speech when speech was shifted down in frequency (equivalent to a lower tape speed), the last bit of the buffer was rejected when sampling of the next buffer commenced. The resultant sound was low-pass filtered to remove any distortion brought about by truncating the replay buffer. Importantly, buffer length was only 10 ms so that when speech was shifted down an octave (only the first half of the buffer used for replay), samples could be out of synchrony by 5 ms maximum, meaning the shifted version was presented virtually in real time. Other features to note about FSF are that the signal level in the shifted version varies with speech level (when speakers produce low intensity sounds, the FSF is also low in intensity, and vice versa). Also, no sound occurs when the speaker is silent (the latter is a feature that is shared with the Edinburgh masker). The two preceding factors limit the noise dose the speaker receives.

The effects on fluency of this (almost real time) ARAI was a marked improvement in fluency in people who stammer even when speakers were instructed to speak at normal rate. Howell, El-Yaniv and Powell's (1987) first study showed that FSF resulted in more fluent speech than DAF or the Edinburgh masker. Later studies have argued that FSF does not produce speech that is superior to DAF speech at short delays (Kalinowski, Armson, Roland-Mieszkowski, Stuart, & Gracco, 1993; Macleod, Kalinowski, Stuart & Armson, 1995). However, these studies have used fast Fourier transform (FFT) techniques to produce frequency shifts. These techniques produce significant delays and the delays are somewhat variable (Howell & Sackin, 2002). Therefore, the studies that claim FSF has the same effect on fluency as DAF have compared FSF plus a short delay, with short-delay DAF. Thus the delay they include under FSF may account for why these studies failed to find a difference between it and DAF whereas Howell et al. (1987) did. (The importance of exact synchrony between altered and recurrent sounds is returned to later where observations about SpeechEasy™ are made.)

A second important point about the Howell, El-Yaniv and Powell (1987) study was that, as mentioned, the effects on fluency were observed even though speakers were told to speak at a normal rate. Therefore, to the extent to which they obeyed instructions, the effects of FSF seem to be independent of rate. This argues against Costello-Ingham's (1993) view that ARAI techniques (DAF in particular) work because they slow overall speech rate. Direct tests of whether fluency-enhancing effects occur when speech rate is varied were made by

Kalinowski et al. (1996) for DAF, and by Hargrave, Kalinowski, Stuart, Armson and Jones (1994), and Natke, Grosser and Kalveram (2001) for FSF. These studies reported that fluency was enhanced whether or not rate was slow (relative to normal speaking conditions). One proviso about the Kalinowski studies is that a global measure of speech rate was taken. It is possible for speakers to speed up global (mean) speech rate while, at the same time, reducing rate locally within an utterance. See Howell and Sackin (2000) for an empirical study that shows fluent speakers display local slowing in singing and local and global slowing under FSF. Also see Howell (in press) for an extended discussion of rate change and its effect on stammering. Until local measures are taken under FSF in people who stammer, it cannot be firmly concluded whether fluency changes are associated with rate change or not, since the speakers might have increased global rate but reduced local rate around the points where disfluencies would have occurred (Howell & Sackin, 2000).

In Howell, El-Yaniv and Powell's (1987) fourth experiment, the effects of presenting FSF just at sound onset (where speakers who stammer have most problems) were compared with those in continuous FSF speech. The effects on fluency did not differ significantly between the two conditions, suggesting that just having FSF at sound onset was as effective as having it on throughout the utterance. This shows that it may be possible to get as much enhancement in fluency when alteration is made just to selected areas in an utterance compared with when alteration is made to the whole utterance. This effect is akin, in some ways, to targeting sections where rate is too high in the Hector aid.

These initial studies suggested that FSF increases fluency and has few secondary effects on speech control (it has little effect on speech rate). Subsequent studies have shown that FSF also has little effect on voice level (it produces a small Fletcher effect rather than a Lombard effect) (Howell, 1990). There is incomplete compensation for shifts in frequency of voice pitch in fluent speakers (Burnett, Senner & Larson, 1997), for upward shifts in speakers who stammer (Natke et al., 2001) and no compensation at all for downward shifts in people who stammer (Natke et al., 2001). Kalinowski's group claims the paucity of secondary effects makes FSF acoustically 'invisible' (and they maintain the same applies to short-duration DAF). They also claim that the minimal changes in speech control under these two forms of ARAI lead speakers to produce fluent, or near fluent, speech (Kalinowski & Dayalu, 2002).

Kalinowski's group has investigated how FSF operates in more natural situations such as over the telephone (Zimmerman, Kalinowski, Stuart, & Rastatter, 1997), or when speakers have to speak in front of audiences (Armson, Foote, Witt, Kalinowski, & Stuart, 1997). They reported that, in both these situations, there are marked improvements in fluency and, therefore, that these procedures may operate in natural environments.

The most recent achievement of the Kalinowski group has been the development of the SpeechEasy™ device which can be worn in the ear and used away from the clinic. This freedom will change the role of the therapist. A move towards delivering therapy outside the clinic has also been taken by those working on the Lidcombe operant therapy (Onslow et al., 1994). It should be noted, however, that application of the Lidcombe program outside the clinic is carefully regulated, the team giving strict guidelines as to what can be done and strictly monitoring that these guidelines are being adhered to.

While Kalinowski and colleagues have stressed how close short delay DAF is to fluent speech, others have noted that even short delays have effects on speech output. For instance, Kalveram and his colleagues at Dusseldorf have established that DAF with short delays, comparable to those used in the SpeechEasy™ device, has effects on the duration of stressed vowels. They report that stressed vowels are prolonged by between 10 and 40% (depending on speech rate and delay) (Kalveram, 2001; Kalveram & Jaencke, 1989).

### ARAI produced by the Speech Easy™ device

Given the rapid introduction and growth in popularity of the Speech Easy™ device, it seems appropriate to take a critical look at the alterations such devices make, and in particular to examine the impact they may have on speech control if they are used in the long term. First, devices that use FFT methods to produce the frequency shift will introduce a timing delay, and this delay may have deleterious effects on speech control, as mentioned above (Novak, 1978). In a technical description of the SpeechEasy™ device (Stuart, Xia, Jiang, Jiang, Kalinowski, & Rastatter, 2003), no details of the temporal delay associated with FSF were given though, based on Howell and Sackin's (2002) observations, these delays may not be negligible. If there are significant delays in the device that carry over into speech when the device is not used, it ought to be redesigned to minimize delay using a speed changing method (such as that used in Howell et al.'s, 1987, original work).

Second, the compression of the speech spectrum by the SpeechEasy™ device, destroys some of the spectral structure when speech is shifted down (Stuart et al., 2003). This would lead to a down-shifted version to be more like noise than the ordinary voice (and possibly an upward-shifted version). This could induce a Lombard effect (increased voice level).

Third, shifting the spectrum shifts the speech formants that carry information about the speech sound spoken. Houde and Jordan (1998) report that long-term exposure to spectrally-shifted speech results in the speaker making compensatory changes so that the speech heard has formants closer to those the speaker intended to produce. The SpeechEasy™ device could also result in vowel quality changes if used in the long term.

The fourth point that should be mentioned is based on the claim of some workers who have disputed whether all speakers have a consistent response to FSF (Ingham, Moglia, Frank, Costello-Ingham & Cordes, 1997). Ingham and colleagues ran two experiments, only the first of which is relevant to the consistency claim. In this study, they tested four subjects under FSF and claimed the effects were not consistent over all their subjects. Though this might raise reservations about general use of FSF there are some procedural details that undermine their statement about the consistency of the FSF effect. Their subject E.S., for instance, reported that “he could speak more easily during the FSF conditions”, but Ingham et al. (1997) did not include him in their second study because they were not able to detect this improvement. The procedure they used was a time-interval procedure on 5-sec long intervals. Virtually all 36 of E.S.'s 5-sec intervals were judged stammered presumably because he had a severe problem), resulting in a ceiling effect with and without FSF (all 36 intervals judged stammered). However, if they had used a shorter interval they would have avoided the ceiling effect and the analysis would probably have resulted in detection of the improvement E.S. reported under FSF (see Howell, Staveley, Sackin, & Rustin, 1998, for further discussion of these and other problems associated with time interval techniques). In fact there are indications with regards to the Ingham et al. paper (from personal reports of their participants and by inspection of the data obtained) that the speech of all four of their speakers improved under FSF. The details of this study do not support the authors' views about whether the effects of FSF are consistent over speakers.

Besides these effects with the frequency shifts created by the SpeechEasy™ device, there are also reasons for supposing that short-delay DAF would affect speech. For instance, the work of Kalveram's group (discussed above) suggests that stressed vowels are lengthened under short-delay DAF.

### Theoretical accounts of DAF and FSF

In this section, two contrasting accounts of why short-delay DAF and FSF produce marked improvement in the fluency of people who stammer and fluent speakers are considered.

Coverage of theories is not, then, comprehensive and, as indicated under 'structure of the review', weighted towards the author's EXPLAN theory. The two theories were selected because they propose that these alterations affect different locations in the central nervous system (CNS). Kalinowski's group maintains that these forms of ARAI operate at high levels in the CNS in speakers who stammer. Howell's group suggest that ARAI operates on low level (probably cerebellar) timekeeping processes in all speakers.

Points made by Kalinowski and co-workers in support of their theory are:

1. DAF at short delays and FSF allows speakers who stammer to produce fluent speech (Kalinowski & Dayalu, 2002). Prolonged speech methods, that also improve fluency (Costello-Ingham, 1993) lead to speech that is not fluent.
2. ARAI works because it presents a second speech signal via perception (Kalinowski & Dayalu, 2002). This second signal creates a situation that is analogous in some ways to choral speech (that is also known to elicit fluent speech). In support of the view that choral speech is fluent, studies have shown that brain image patterns of people who stammer under choral speaking conditions are almost indistinguishable from fluent speakers' patterns.
3. The central mechanism that is affected by ARAI is one that links production with perception (the mirror neuron system, Kalinowski & Saltuklaroglu, 2003). Mirror neurons discharge when an action is either performed or is observed (i.e. motor and sensory properties coexist in the same neuron). Mirror neurons could affect fluency, as they are found in Broca's speech motor area (Nishitani & Hari, 2000).
4. The mirror neuron system is, according to Kalinowski, important in early development (children's imitations). It appears to be used less as speakers get older. However, the second signal under ARAI activates the mirror neuron system. This assists production and allows fluency to be regained in speakers who stammer.
5. The changes in fluency in people who stammer occur passively when ARAI is presented (Saltuklaroglu, Dayalu & Kalinowski, 2002) and these passive changes occur because the central mirror neuron system is affected directly. This contrasts with the changes that arise with techniques like prolonged speech that requires the speaker to make an active change. Such active changes can eventually affect the same system that passive changes influence. This could account for cases where speakers who stammer are successfully treated by techniques like prolonged speech.

Several observations are now made about points 1) – 5):

1. The hypothesis that speech under ARAI produces fluent speech predicts that there will be no differences between fluent material and ARAI material. Statistically speaking, this is a situation where the null hypothesis is predicted which is against a fundamental principle of statistics. The work of Kalinowski's group actually establishes that ARAI leads to high levels of stutter-free speech that, it is claimed, sounds natural. Even though ARAI speech is closer to the speech produced by fluent speakers than the end-product of prolonged speech regimes, it still may not be fluent as the studies in the previous section indicate. Also, methods of measuring various aspects of voice control are constantly being improved and these improved measures may reveal important, yet subtle, effects on fluency. For example Kalveram and colleagues duration measurements of stressed vowels has found effects of short-delay DAF. There are reasons for supposing that speakers may change the position of the articulators when FSF is delivered (Houde & Jordan, 1998). It can be inferred from Houde and Jordan's (1998) study that such changes



in articulator position are subtle and would not be easily detectable by perceptual assessment alone. An appropriate technique for establishing whether these occur would be formant frequency analysis and no such studies have been reported to date on stammered speech after exposure to FSF. Generally speaking, these two examples illustrate that there are grounds for considering that differences in fluency between ARAI and fluent speech that are hard to detect using simple measures may be detectable when improved techniques are employed.

2. There are 'second signals' (to use Kalinowski's term) that affect the fluency of people who stammer that are not speech. One example, discussed above, is Howell and El-Yaniv's metronome signal where a click is triggered by the speaker's speech (not at a pre-set pace). It is hard to imagine how this signal could be used by the mirror neuron system as it bears no relation at all to speech, yet it improves the fluency of people who stammer.
3. The Howell and Archer (1984) study on fluent speakers showed that the effects of ARAI arise at a lower level in the CNS than mechanisms involved in speech perception. This would show that central perceptual processes are not involved in the case of DAF, assuming Howell and Archer's (1984) result applies to people who stammer under DAF as well as to fluent speakers. Other problems for accounts that maintain that ARAI influences central levels involved in perception have been extensively discussed recently (Howell, 2002; Howell, in press).
4. To work, the mirror neuron system has to have some input from perception to reflect into production at the time the speech is being produced. However, as indicated earlier, commercially available ARAI techniques that improve the fluency of people who stammer produce perceptual information after production is required. Thus, there is an inherent delay between production of a sound and when the altered sound is received with DAF; the Edinburgh masker has a lag too and there are grounds for supposing that this also applies to the SpeechEasy™ device. It is, of course, possible to modify the mirror neuron concept. For instance, the mirror neurons could be made more flexible both in terms of, a) how closely timed speech events and the perceptual events they give rise to need to be, and b) how similar the perceptual events need to be relative to the linguistic events they reflect. Though it is appropriate to postulate such flexibility, neurological data would be needed to support such temporal and linguistic flexibility before they are taken as fact. Finally, endowing mirror neurons with too much flexibility seems inadvisable. There needs to be some delimitation of the range of what perceptual events trigger activity in these neurons otherwise they lose their selectivity in linking actions with the perceptual events that gave rise to them.
5. ARAI is supposed to affect the mirror neuron system directly. Techniques that train speakers to relearn motor patterns, operate at the motor level initially and, only when the patterns have been established, can they be transmitted to the mirror neuron system. Kalinowski's group proposes, these techniques then affect this system in a similar way to ARAI. Therefore, ARAI and learning techniques operate initially on different mechanisms (ARAI affects speech "passively", bypassing the peripheral level). An implication of this position is that there is no single factor that explains both how ARAI and motor processes affect fluency (underlined by their dismissal of Costello-Ingham's, 1993, proposal that rate underlies ARAI and prolonged speech procedures). Consistent with the Kalinowski group's view, there do seem to be grounds for considering that the time courses of ARAI and operant procedures differ (e.g. the Lidcombe program). ARAI affects fluency 1) in the short term, and 2) these effects are restricted mainly to the periods during which ARAI is presented. In contrast, the Lidcombe program 1) does not have dramatic effects

short term, but 2) the effects on fluency are reported to be maintained for longer (and in some cases result in fluency being permanently regained). However, though the different timecourses of the effects would be consistent with the two procedures affecting different CNS locations, the proposal that ARAI (just central) works in a different way to operant procedures (peripheral and central) is not parsimonious.

Howell and co-workers' EXPLAN model has been reviewed extensively in recent publications (Howell, 2002, in press; Howell & Au-Yeung, 2002). It is a general model of spontaneous speech control that attempts to explain: 1) developmental changes in patterns of stammering, and 2) how stammering relates to fluent speech, as well as 3) the effects of ARAI. Detailed review of the first two topics is beyond the scope of this article, but some background information is necessary. The basic idea behind the EXPLAN model is that cognitive-linguistic planning (PLAN) processes are independent of motor execution (EX) processes. The role of the planning processes is to supply a plan for an utterance when the motor execution processes have finished producing the previous utterance. Disfluencies arise when the plan is not ready at this time. In a phrase like "I split it", the comparatively complex word "split" is likely to be the one that is not ready in time for execution. If this is the case speakers may do one of two things: First, they may repeat or hesitate on the prior word (producing, for example, "I, I, split it"). Howell (in press) refers to these events as stalling disfluencies. Second, since plans are assumed to be generated left to right, speakers can commence "split" using the plan for the first part of the word which is available. Planning continues while this first part is being uttered, as this process is independent of execution. The remainder of the plan may be generated in the time taken to execute the first part. However, the plan can run out and result in disfluencies involving just the first part of the word (e.g. "sssplits", "s.s.split"). Howell (in press) refers to these as advancing stutterings. The latter are characteristic features of adult stammered speech in a variety of languages (Au-Yeung, Vallejo Gomez & Howell, 2003; Dworzynski, Howell, Au-Yeung & Rommel, in press; Howell, Au-Yeung & Sackin, 1999).

This account implies that the adult pattern of stammering is a result of attempting to produce speech locally at too fast a rate. EXPLAN proposes that this pattern can be avoided in two ways. First, speakers can change speech execution rate using a timekeeper that changes execution rate directly (Howell, 2002). Second, speakers can change the way the chaining process between planning and execution operates without involving the timekeeper (Howell, in press). Stallings and advancements are different ways of changing the operation of the chaining between planning and execution processes when the plan for the following word is not ready. Stalling repeats a plan (uses a pre-existing plan) or interrupts speech to gain more time and does not involve the problem word at all. This option is frequently used by fluent speakers (Howell, Au-Yeung & Sackin, 1999), so it does not have deleterious effects on long-term fluency. Advancing gambles that execution time is long enough to generate the remainder of the plan. Advancing is problematic as it can fail (as indicated by the fact that it can lead to disfluencies on part of a word). Though the mechanisms involved differ, both execution rate and one of the two ways of changing the chaining between planning and execution are, generically speaking, ways of changing speech rate.

EXPLAN contrasts with Kalinowski's account on all five of the points outlined above. The contrasts, and data that support the EXPLAN view, are as follows:

1. ARAI produces fluent speech by affecting a timekeeping process that controls execution rate directly. Other ways of affecting timing (whether using the timekeeping mechanism or not) improve fluency by gaining extra planning time for 'problem' words. According to this principle, learning procedures control rate in the planning-execution chain. These procedures alert speakers to situations where they are adopting a maladaptive way of dealing with speech when its plan is not

complete. The different mechanisms and mode of achieving rate control involved with operant procedures could explain why they take longer to affect fluency than ARAI. Though operant procedures take longer, the way they achieve fluency is the same as in ARAI.

2. ARAI does not so much produce a “second signal” that is speech, as introduce a second rhythmic signal. This second rhythmic signal affects speech control (particularly when it is slightly out of synchrony), by changing operation of the timekeeper. See Howell and Sackin (2002) for evidence that supports the view that DAF affects a timekeeping process in the cerebellum.
3. ARAI is not effective because it affects a central process that links speech perception and production. Many ARAI manipulations that affect fluency of people who stammer are not speech sounds. Examples include Howell and Archer's (1984) noise stimulus, Howell and El-Yaniv's (1987) metronome signal positioned at syllable onset, and even a flashing light (Kuniszuk-Jozkowiak, Smolka & Adamczyk, 1996).
4. Synchronous and delayed asynchronous signals all affect operation of the timekeeper (Howell & Sackin, 2002). A speech signal is not needed in advance to prime production (Howell & Archer, 1984). The EXPLAN process does not fail because perceptual information is not available prior to, or even during, production (as required by the mirror-neuron system).
5. Rate control takes place (albeit in different ways) for ARAI and motor-learning procedures (see Howell, in press, for a detailed description of how the two interrelate). Possibilities are opened up (see the next section), given that these two procedures have the common basis of gaining planning time.

The theory of Kalinowski's group and EXPLAN were selected as contrasting views about what level of speech control is affected by ARAI. Other theories in the area either do not include accounts of the fluency-enhancing effects of FSF (Neilson & Neilson, 1991) or maintain that there are influences at both peripheral and central areas of the central nervous system (Kalveram, 2001; Kalveram & Jaencke, 1989). Both these have similarities and differences with respect to EXPLAN. The similarities in Kalveram's model, for example, concern the planning phase for serialisation of speech units (words, syllables, phonemes) that must be prepared in advance of motor execution. A dissimilarity concerns whether speakers use acoustic-phonetic information in the control of speaking (Dusseldorf group), or whether the control system crashes until timing recovers if planning and execution do not match (EXPLAN).

## Summary and future possibilities

The fluency-enhancing effects of ARAI are indisputable. Short delay DAF and synchronous alterations (FSF) produce speech that sounds very nearly fluent. Devices like SpeechEasy™ have obvious attractions to a person who stammers because they produce at least temporary fluency. The main question to be addressed here is whether the aid ought to be used continuously or intermittently (grounds are given for supposing that intermittent presentation might promote carry-over of fluent patterns). Before that question is addressed, it should be noted in passing that even if the device only works while speech is altered continuously (i.e. there is no carry-over of the fluency-enhancing effects), it would still be useful (over the phone, with an audience or in other situations the owner chose to use it).

My group's theoretical perspective (EXPLAN) suggests that rate control lies behind the effectiveness of these devices. However, dramatic slowing (as with prolonged speech techniques) is unnecessary; slowing only needs to occur in the local vicinity of a difficult

word. Also, having ARAI on all the time might not promote transfer of the fluent behavior induced. As stammering occurs intermittently throughout speech, 'rate' (understood in the general sense used earlier) only needs to be altered in the vicinity of these episodes. This suggests that ARAI ought to be targeted only on or around problematic sounds. Targeting particular episodes in a similar way is a feature of operant treatment procedures.

Looked at from the point of view of continuous delivery of ARAI sounds, it does not appear to be sensible to present these alterations on episodes within a stammerer's speech which are fluent, for several reasons. Transfer would not be promoted. It is not certain that FSF and short-delay DAF produce absolutely fluent speech, and these residual nonfluent behaviors could be transferred to post-treatment speech (Novak, 1978). There may be long-term effects of FSF (Houde & Jordan, 1998) not evident in the current short-term studies that impact on long-term fluency. Any procedure that restricts exposure to ARAI while at the same time maintaining high rates of fluency may be advantageous (see the above discussion of the Hector aid and Howell et al., 1987, experiment 4).

Targeting disfluencies for a dose of ARAI also opens up possibilities that allow effects (known in the animal operant literature) that should produce maintenance of fluent behaviors, to be exploited. A partial reinforcement schedule retains response behaviors for longer than responses that are continuously reinforced. If techniques were available that allowed regions that contain disfluent episodes to be targeted for ARAI, schedules of reinforcement could be manipulated to see whether this applies to part-presentation of ARAI. Though ARAI and operant procedures have been used jointly in treatments, to date there has been no study that administers ARAI on a partial reinforcement schedule. One reason for this may be that training under partial reinforcement protocols takes a long time. Nevertheless, until such studies have been completed, the possibility that ARAI could lead to long-term recovery cannot be ruled out. One possible way that alterations could be targeted on regions that are disfluent (or are at high risk of being so) would be to use speech rate as in the pioneering work on the Hector aid.

## Acknowledgments

This research was supported by the Wellcome Trust. Thanks to Professor Kalveram, Dr Kalinowski, and Messrs. Dayalu and Saltuklaroglu for their help with processing of this manuscript.

## References

- Albright MAH, Malone JY. The relationship of hearing acuity to stammering. *Exceptional Children*. 1942; 8:186–190.
- Armson J, Foote S, Witt C, Kalinowski J, Stuart A. Effect of frequency altered feedback and audience size on stuttering. *European Journal of Disorders of Communication*. 1997; 32:359–366. [PubMed: 9474287]
- Au-Yeung J, Vallejo Gomez I, Howell P. Exchange of disfluency from function words to content words with age in Spanish speakers who stutter. *Journal of Speech, Language and Hearing Research*. 2003; 46:754–765. [PubMed: 14697001]
- Backus O. Incidence of stuttering among the deaf. *Annals of Otology, Rhinology and Laryngology*. 1939; 47:632–635.
- Black JW. The effect of delayed sidetone upon vocal rate and intensity. *Journal of Speech and Hearing Disorders*. 1951; 16:56–60.
- Borden GJ. An interpretation of research on feedback interruption in speech. *Brain & Language*. 1979; 7:307–319. [PubMed: 455050]
- Brown, GS.; Campbell, DP. *Principles of servomechanisms*. New York: Wiley; 1948.
- Burnett TA, Senner JE, Larson CR. Voice F0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice*. 1997; 11:202–211. [PubMed: 9181544]

- CCITT. Interactions between sidetone and echo. CCITT – International Telegraph and Telephone Consultative Committee; 1989a. Contribution, com XII, no BB
- CCITT. Experiments on short-term delay and echo in conversation. CCITT – International Telegraph and Telephone Consultative Committee; 1989b. Contribution, com XII, no AA
- Chase RA, Sutton S, Rapin I. Sensory feedback influences on motor performance. *Journal of Auditory Research*. 1961; 1:212–223.
- Cherry C, Sayers B. Experiments upon the total inhibition of stammering by external control and some clinical results. *Journal of Psychosomatic Research*. 1956; 1:233–246. [PubMed: 13377328]
- Costello-Ingham JC. Current status of stuttering and behavior modification – 1. Recent trends in the application of behavior application in children and adults. *Journal of Fluency Disorders*. 1993; 18:27–44.
- Curlee RF, Perkins WH. Conversational rate control for stuttering. *Journal of Speech and Hearing Disorders*. 1969; 34:245–250. [PubMed: 5821127]
- Dewar A, Dewar AW, Austin WTS, Brash HM. The long-term use of an automatically triggered auditory feedback masking device in the treatment of stammering. *British Journal of Disorders of Communication*. 1979; 14:219–229.
- Donath TM, Natke U, Kalveram KT. Effects of frequency-shifted auditory feedback in voice F<sub>0</sub> contours in syllables. *Journal of the Acoustical Society of America*. 2002; 111:357–366. [PubMed: 11831808]
- Dworzynski K, Howell P, Au-Yeung J, Rommel D. Stuttering on function and content words across age groups of German speakers who stutter. *Journal of Multilingual Communication Disorders*. in press.
- Fairbanks G. Selected vocal effects of delayed auditory feedback. *Journal of Speech and Hearing Disorders*. 1955; 20:333–345. [PubMed: 13272227]
- Fletcher, H.; Raff, GM.; Parmley, F. Study of the effects of different sidetones in the telephone set. Western Electrical Company; 1918. Report no. 19412, Case no. 120622
- Goldiamond, I. Stuttering and fluency as manipulatable operant response classes. In: Krasner, L.; Ullman, L., editors. *Research in behavior modification*. New York: Holt, Rhinehart and Winston; 1965. p. 106-156.
- Guenther, F. Neural modeling of speech production. In: Maassen, B.; Hulstijn, W.; Kent, R.; Peters, HFM.; van Lieshout, PHMM., editors. *Speech motor control in normal and disordered speech*. Nijmegen: Uitgeverij Vantilt; 2001. p. 12-15.
- Ham R, Steer MD. Certain effects of alterations in auditory feedback. *Folia Phoniatica*. 1967; 19:53–62. [PubMed: 5584824]
- Hargrave S, Kalinowski J, Stuart A, Armson J, Jones K. Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech and Hearing Research*. 1994; 37:1313–1319. [PubMed: 7877290]
- Houde J, Jordan MI. Adaptation in speech production. *Science*. 1998; 279:1213–1216. [PubMed: 9469813]
- Howell P. Changes in voice level caused by several forms of altered feedback in normal speakers and stutterers. *Language and Speech*. 1990; 33:325–338. [PubMed: 2133911]
- Howell, P. The EXPLAN theory of fluency control applied to the treatment of stuttering by altered feedback and operant procedures. In: Fave, E., editor. *Pathology and therapy of speech disorders*. Amsterdam: John Benjamins; 2002.
- Howell P. Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communicative Sciences and Disorders*. in press.
- Howell P, Archer A. Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics*. 1984; 36:296–302. [PubMed: 6522222]
- Howell, P.; Au-Yeung, J. The EXPLAN theory of fluency control and the diagnosis of stuttering. In: Fava, E., editor. *Current Issues in Linguistic Theory series: Pathology and therapy of speech disorders*. Amsterdam: John Benjamins; 2002. p. 75-94.
- Howell P, Au-Yeung J, Pilgrim L. Utterance rate and linguistic properties as determinants of speech dysfluency in children who stutter. *Journal of the Acoustical Society of America*. 1999; 105:481–490. [PubMed: 9921672]



- Howell P, Au-Yeung J, Sackin S. Exchange of stuttering from function words to content words with age. *Journal of Speech, Language and Hearing Research*. 1999; 42:345–354. [PubMed: 10229451]
- Howell P, El-Yaniv N. The effects of presenting a click in syllable-initial position on the speech of stutterers: Comparison with a metronome click. *Journal of Fluency Disorders*. 1987; 12:249–256.
- Howell, P.; El-Yaniv, N.; Powell, DJ. Factors affecting fluency in stutterers when speaking under altered auditory feedback. In: Peters, H.; Hulstijn, W., editors. *Speech Motor Dynamics in Stuttering*. New York: Springer Press; 1987. p. 361-369.
- Howell P, Marchbanks RJ, El-Yaniv N. Middle ear muscle activity during vocalisation in normal speakers and stutterers. *Acta Oto-Laryngologica*. 1986; 102:396–402. [PubMed: 3788538]
- Howell P, Powell DJ. Hearing your voice through bone and air: Implications for explanations of stuttering behaviour from studies of normal speakers. *Journal of Fluency Disorders*. 1984; 9:247–264.
- 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; 9:772–784.
- Howell P, Sackin S. Speech rate manipulation and its effects on fluency reversal in children who stutter. *Journal of Developmental and Physical Disabilities*. 2000; 12:291–315. [PubMed: 18259598]
- Howell P, Sackin S. Timing interference to speech in altered listening conditions. *Journal of the Acoustical Society of America*. 2002; 111:2842–2852. [PubMed: 12083218]
- Howell P, Staveley A, Sackin S, Rustin L. Methods of interval selection, presence of noise and their effects on detectability of repetitions and prolongations. *Journal of the Acoustical Society of America*. 1998; 104:3558–3567. [PubMed: 9857514]
- Howell, P.; Wingfield, T.; Johnson, M. Characteristics of the speech of stutterers during normal and altered auditory feedback. In: Ainsworth, WA.; Holmes, JN., editors. *Proceedings Speech 88; 7th Federation of Acoustical Societies of Europe conference; Edinburgh*. 1988; Edinburgh: Institute of Acoustics; 1988. p. 1069-1076. Vol. 3
- Ingham, RJ.; Fox, PT.; Ingham, JC. An H<sub>2</sub><sup>15</sup>O positron emission tomography (PET) study on adults who stutter. In: Hulstijn, W.; Peters, HFM.; van Lieshout, PHHM., editors. *Speech Production: Motor Control, Brain Research and Fluency Disorders*. Amsterdam: Elsevier; 1997. p. 293-305.
- Ingham RJ, Moglia RA, Frank P, Costello-Ingham J, Cordes A. Experimental investigation of the effects of frequency-altered feedback on the speech of adults who stutter. *Journal of Speech, Language and Hearing Research*. 1997; 40:361–372. [PubMed: 9130204]
- Kalinowski J, Armson J, Roland-Mieszkowski M, Stuart A, Gracco V. Effects of alterations in auditory feedback and speech rate on stuttering frequency. *Language and Speech*. 1993; 36:1–16. [PubMed: 8345771]
- Kalinowski J, Dayalu V. A common element in the immediate inducement of effortless, natural-sounding, fluent speech in stutterers: “The Second Speech Signal”. *Medical Hypotheses*. 2002; 58:61–66. [PubMed: 11863399]
- Kalinowski J, Saltuklaroglu T. Speaking with a mirror: Engagement of mirror neurons via choral speech and its derivatives induces stuttering inhibition. *Medical Hypotheses*. 2003; 60:538–43. [PubMed: 12615517]
- Kalinowski J, Stuart A, Sark S, Armson J. Stuttering amelioration at various auditory feedback delays and speech rates. *European Journal of Disorders of Communication*. 1996; 31:259–269. [PubMed: 8944848]
- Kalveram, K. Th. Neurobiology of speaking and stuttering. In: Bosshardt, HG.; Yaruss, JS.; Peters, HFM., editors. *Fluency Disorders: Theory, Research, Treatment and Self-help*. Proceedings of the Third World Congress of Fluency Disorders; Nijmegen: Nijmegen University Press; 2001. p. 59-65.
- Kalveram, K.Th.; Jaencke, L. Vowel duration and voice onset time for stressed and nonstressed syllables in stutterers under delayed auditory feedback condition. *Folia Phoniatrica*. 1989; 41:30–42. [PubMed: 2759514]
- Kawato M, Furakawa K, Suzuki R. A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*. 1987; 57:169–185. [PubMed: 3676355]

- Kern SH. Der einfluss des horens auf das stottern. *Archiv Psychiatrischer Nervenkrankheiten*. 1932; 97:428–450.
- Kuniszuk-Jozkowiak W, Smolka E, Adamczyk B. Effect of acoustical, visual and tactile reverberation on speech fluency of stutterers. *Folia Phoniatica & Logopedics*. 1996; 48:193–200.
- Lane HL, Tranel B. The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*. 1971; 14:677–709.
- Lee BS. Effects of delayed speech feedback. *Journal of the Acoustical Society of America*. 1950; 22:824–826.
- Lee BS. Artificial stutter. *Journal of Speech and Hearing Disorders*. 1951; 15:53–55.
- Lombard E. Le signe de l'elevation de la voix. *Annales des Maladies de l'Oreille, du Larynx, du Nez et du Pharynx*. 1911; 37:101–119.
- Lotzmann G. Zur Anwedung variiertes verzogerungszeiten bei balbuties. *Folia Phoniatica & Logopedics*. 1961; 13:276–312.
- Macleod J, Kalinowski J, Stuart A, Armson J. Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. *Journal of Communication Disorders*. 1995; 28:217–228. [PubMed: 8530718]
- Natke U, Grosser J, Kalveram KT. Fluency, fundamental frequency, and speech rate under frequency shifted auditory feedback in stuttering and nonstuttering persons. *Journal of Fluency Disorders*. 2001; 26:227–241.
- Neelley JN. A study of the speech behaviors of stutterers and nonstutterers under normal and delayed auditory feedback. *Journal of Speech and Hearing Disorders Monograph*. 1961; 7:63–82.
- Neilson, MD.; Neilson, PD. Adaptive model theory of speech motor control and stuttering. In: Peters, HFM.; Hulstijn, W.; Starkweather, CW., editors. *Speech motor control and stuttering*. Amsterdam: Excerpta Medica; 1991. p. 149-156.
- Nessel E. Die verzogerte Sprachruckkopplung (Lee Effect) bei Stotteren. *Folia Phoniatica*. 1958; 10:199–204. [PubMed: 13619922]
- Nishitani N, Hari R. Temporal dynamics of cortical representation for action. *Proceedings on the National Academy of Sciences of the United States of America*. 2000; 97:913–918. [PubMed: 10639179]
- Novak A. The influence of delayed auditory feedback in stutterers. *Folia Phoniatica*. 1978; 30:278–285. [PubMed: 730093]
- Onslow M, Andrews C, Lincoln M. A control/experimental trial of operant treatment for early stuttering. *Journal of Speech and Hearing Research*. 1994; 37:1244–1259. [PubMed: 7877284]
- Ryan, B. *Programmed stuttering therapy for children and adults*. Springfield: C. C. Thomas; 1974.
- Saltuklaroglu T, Dayalu VN, Kalinowski J. Reduction of stuttering: The dual inhibition hypothesis. *Medical Hypotheses*. 2002; 58:67–71. [PubMed: 11863400]
- Shearer WM. Speech behavior of middle ear muscles during stuttering. *Science*. 1966; 152:1280. [PubMed: 5937122]
- Stuart A, Xia S, Jiang Y, Jiang T, Kalinowski J, Rastatter MP. Self-contained in-the-ear device to deliver altered auditory feedback: Applications for stuttering. *Annals of Biomedical Engineering*. 2003; 31:233–7. [PubMed: 12627830]
- Soderberg GA. A study of the effects of delayed side-tone on four aspects of stutterers' speech during oral reading and spontaneous speech. *Speech Monographs*. 1960; 27:252–253.
- Sutton S, Chase RA. White noise and stuttering. *Journal of Speech and Hearing Research*. 1961; 4:72.
- Teig E. Differential effect of graded contraction of middle ear muscles on the sound transmission of the ear. *Acta Physiologica Scandinavica*. 1973; 88:387–391. [PubMed: 4751174]
- Webster RL, Lubker B. Masking of auditory feedback in stutterers' speech. *Journal of Speech and Hearing Research*. 1968a; 11:221–222. [PubMed: 5648537]
- Webster RL, Lubker B. Interrelationships among fluency-producing variables in stuttered speech. *Journal of Speech and Hearing Research*. 1968b; 11:754–766. [PubMed: 5719230]
- Wingate, M. *Foundations of stuttering*. New York: Academic Press; 2002.

Zimmerman S, Kalinowski J, Stuart A, Rastatter MP. Effect of altered auditory feedback on people who stutter during scripted telephone conversations. *Journal of Speech, Language, and Hearing Research*. 1997; 40:1130–1134.

**Table 1**

summarizes the important differences between the proposal of Kalinowski's group and EXPLAN.

<b>Level</b>	<b>Kalinowski</b>	<b>EXPLAN</b>
Linguistic	AAF	
Linguistic- motor		Operant
Motor	Motor learning (e.g. operant procedures)	ARAI