The handle http://hdl.handle.net/1887/19818 holds various files of this Leiden University dissertation.

**Author:** Umanski, Daniil  
**Title:** The use of technological support in communication disorders: how development of computer-based tools can refine the treatment of motor speech disorders  
**Date:** 2012-09-18
He who does not understand your silence will probably not understand your words.

_Elbert Hubbard, 1911_
6.1 Introduction

In this chapter, an improvement to the current methodologies of altered auditory feedback (AAF) for people who stutter is proposed. An attempt is made to examine adaptive feedback procedures, which are based on real-time analysis of the speech signal, and the dynamic activation of auditory feedback. Specifically, we propose to deal with the difficulty of initiating speech at phrase onsets by selectively providing a feedback signal only in the silent intervals between phrases. In order to evaluate the proposed AAF procedure, adults who stutter produced spontaneous speech under experimental feedback conditions.

Adaptive auditory feedback procedures

Altered auditory feedback (AAF) is known to reduce stuttering by 50-80% in some people who stutter (Lincoln et al., 2006). AAF can be defined as a manipulation of one’s speech signal, in which the altered signal is fed back to the speaker throughout the act of speech. Two most common forms of AAF are delayed auditory feedback (DAF), whereby speakers hear their own voice with a short time delay, and frequency altered feedback (FAF), whereby the frequency spectrum of the speaker’s voice is shifted up or down (for background information on the theoretical and methodological aspects of AAF procedures, please see Chapter 2.4).

One obvious limitation of AAF procedures is that they provide no signal at the moment when phrases are initiated. This is due to the fact that current AAF methods operate on the basis of a static feedback function. In other words, the feedback signal is given continuously, at the same output level, with the same delay/shift settings, regardless of any property of the audio signal produced by the speaker. One way to address these limitations is to design adaptive rather than static feedback functions, with the idea of selectively targeting those regions of speech which are at higher risk of being dysfluent (Howell, 2004). In his extensive overview of altered auditory feedback methodology, Howell suggests that AAF signals should be targeted only on, or around problematic locations. He further proposes that procedures which restrict the exposure to AAF signals while maintaining higher levels of fluency may be advantageous.

In general, an adaptive feedback procedure operates by analysing some property of the speech signal in real-time, and dynamically modifying the output audio signal in relation to that property. The analysis process thus aims to detect regions in the speech stream which correspond to certain ‘states’ of the speaker. In the case of stuttering, problematic regions include phrase onsets, regions in which a block is suspected, or regions in which speech rate is too fast (Wingate and Howell, 2002). Once a target region is detected, an external stimulus signal is then activated to assist speakers with maintaining their flow of speech. Once the system detects that the speaker shifted to another ‘state’, which is not
Adaptive AAF procedures for adults who stutter

considered vulnerable for dysfluencies, the external signal can be attenuated, or withdrawn altogether. It is important to note that the described framework does not involve attempting to detect dysfluent events per se, but rather aims at predicting problematic areas in a speech stream.

Targeting speech rate within a scheme of adaptive feedback is not a new idea. As a classic example, Howell mentions the ‘Hector aid’, a device pioneered by British engineers more than 30 years ago. The device measured speech rate using an audio input, and a vibrator switched on if speech rate was outside an acceptable range, in order to signal the speaker to slow down. The rationale for targeting speech rate modulation comes from evidence that dysfluencies are more likely to occur in the fast-rate areas of speech (Smith and Kleinow, 2000). Although a detailed report about the ‘Hector aid’ is unavailable, the idea behind this device is a good exemplification of the approach advocated in this chapter.

Providing an auditory bridge

The adaptive feedback procedure explored in this study addresses the limitation of current AAF procedures with regard to phrase initiation. A variety of experiments demonstrated the effect of priming fluent speech initiation by means of external auditory cues, with the help of metronome beats, AAF signals, choral speech, and other auditory cues (Azrin et al., 1968; Kuniszyk-Jó kowiak et al., 1997; Howell and El-Yaniv, 1987; Natke et al., 2001; Alm, 2004). Therefore, we assume that an auditory cue present at the moment of phrase initiation has the potential of priming a speaker who stutters to initiate a phrase fluently. Speaking of a ‘phrase’, we refer to a continuous interval of speech, bounded by a silent pause with a duration of 200 milliseconds, or more (Fant et al., 2003). This definition may only partially correspond to the one of the intonation phrase, since the relation between intonation phrase boundaries and silent pauses is not entirely agreed upon (Bulyko and Östendorf, 2001; Yoon et al., 2007; Wagner, 2010). For example, a study by Yoon et al. (2007) revealed that less than 50% of phrase boundaries in English are followed by a silent interval, suggesting that silent pauses alone may not be systematically used as a cue to phrase boundary presence, and that listeners rely on additional acoustic cues, such as pre-boundary lengthening (Lin and Fon, 2009).

In order to activate an audio cue just before the onset of an upcoming phrase, a temporal reference point is needed, which can be detected in the speech stream, allowing to predict that a phrase onset is about to take place. The detection of a silence interval following the termination of the previous phrase is a way to anticipate the timing of the subsequent phrase initiation. Therefore, in our adaptive procedure, a signal is activated only during the silent intervals between phrases in order to prime the initiation of subsequent phrases. We propose that this procedure would provide a kind of an ‘auditory bridge’ between phrases,
6.2 Methods

Participants

The study included 12 adults who stutter (M = 32.3 years, S.D. = 11.5). Eleven were male and one was female. Participants were recruited by announcements placed in clinics for stuttering therapy throughout The Netherlands. All participants reported stuttering onset in childhood, none had severe hearing loss, and all were native Dutch speakers. Ten participants had no previous experience with AAF, and two had a short experience over 3 years prior to the study.
Procedure

The experiments took place in a quite room. Participants were seated approximately 1.5 meters in front of the investigator and a video camera. Participants were asked to produce a monologue under four auditory feedback conditions (Figure 6.1 provides a schematic illustration of the experimental conditions):

(a) no-feedback condition – participants were fitted with headphones, but did not receive any auditory signal while speaking, and could hear their own voice.

(b) a standard DAF condition – the audio signal coming into the microphone was fed back to the participants with a delay of 100 ms. This delay duration was chosen as an average, standard setting used in AAF research, as well as in AAF devices (Lincoln et al., 2006).

(c) an adaptive feedback condition – an audio signal was delivered only during the detected silent intervals. Upon silence detection, an audio buffer containing the most recent 1500 ms is played back until a new phrase onset is detected, with maximum duration of 1500 ms. A speech signal was chosen as the external auditory cue in order to remain in the same auditory domain as the standard DAF procedure, so that comparisons could be made. The choice of the buffer duration is outlined in refbridge.

(d) a combination of conditions (b) and (c) – since part of dysfluencies may occur at mid-phrase positions as well (Saltuklaroglu et al., 2009), the adaptive feedback condition (c) was supplemented with an additional delayed signal (100 ms) during detected speech episodes. During detected speech episodes, the audio signal was fed back to the participants with a delay of 100 ms. Upon a detected silent segment, an audio buffer containing the most recent 1500 ms is played back until a new phrase onset is detected, with maximum duration of 1500 ms.

Figure 6.1: schematic representations of conditions (b) - left, (c) - middle, and (d) - right. The top signal track represents the speech output of the participant, while the bottom track represents the audio signal received by the participant through headphones.
6.2. Methods

The order of presenting the various conditions was randomized across participants. For each condition, the investigator presented a new topic for discussion, asking the participant for his/her opinion on the topic, and eventually signaled the end of discussion once enough speech material was collected (roughly 3 minutes per condition). The same set of questions and topics was used for all participants. A monologue task was chosen since it is a variant of spontaneous speech, which realistically approximates everyday communication situations. Participants were asked to speak in their normal, natural manner, with a normal speech rate, and as far as possible, to avoid using any fluency inducing techniques. In order to prevent the investigator’s voice being fed back to the participant, the investigator spoke only when it was necessary to re-engage the participant in the monologue.

After each condition, the investigator asked the participant to reflect on the experience with the current auditory feedback procedure, as well as to rate the relative level of comfort with the audio signal. A 10-point visual analogue scale was presented on a computer screen, and the participant indicated a score on that scale. This break served to avoid possible carry-over effects between conditions, but also provided valuable insights into the way participants experienced their speech in the given feedback mode. At the end of the experiment, participants were asked to give their general impressions and personal comments on the methods and ideas explored in this study.

Apparatus

All experiment sessions were recorded through a video camera (Panasonic S150). The audio was digitized in real-time by the ZOOM4H microphone and an external audio card. The DAF function was implemented within the MAX/MSP environment (www.cycling74.com). In order to implement the silence/onset detection required for the adaptive feedback procedure, the following real-time signal processing scheme was employed:

1. The raw linear amplitude was converted to amplitude in decibels, so that $0 \text{dB} = 1$ (full amplitude). Treating amplitude on a logarithmic scale provides a better approximation for the perceptual magnitude of changes in intensity.

2. Input level in decibels was then expanded by considering a noise-floor threshold (set manually by the investigator to set the ambient noise inside the experiment room). The expanded level was calculated as a percentage of the above-threshold range.$^2$

$^1$Amplitude expressed in decibels, with: $dB = 20 \times \log_{10}(A)$

$^2$Noise floor canceling is an important step in audio processing, as it can help avoid irrelevant fluctuations of the signal, and make the detection function more robust.
3. Next, the expanded signal was compared against a silence/onset threshold, which was set at 15% of the full dB range (and manually adjusted if necessary). Upon a threshold crossing in either direction (from ‘speech’ to ‘silence’, or the other way around) the system waited for 200 milliseconds before issuing a silence/onset trigger. If, however, a crossing occurs in the opposite direction during these 200 milliseconds, the timer is reset, and no trigger is issued. This mechanism provides a low-pass filter for the detection function and prevents spurious silence/onset triggers.

**Data analysis**

For each speech sample, the fraction of dysfluent phrase onsets from the total number of phrase onsets was calculated. Phrase onsets were defined as moments of speech initiation after an interval of silence longer than 400 ms (Swerts and Geluykens, 1993). A phrase onset was defined as dysfluent if a stuttering-like dysfluency has occurred on, or before the first stressed syllable (Natke et al., 2002). Since the adaptive feedback procedure targets to enhance the initiation of phrases, the proportion of dysfluent phrase onsets reflects how well speakers deal with initiating phrases throughout their speech.

Further, in order to evaluate the overall fluency levels across the experimental conditions, the percentage of discontinuous speech time (PDST) was calculated. In order to arrive at PDST scores, a two-step analysis of the original raw audio samples was performed. First, all segments from the sample are removed which do not belong to the speech of the participant. This includes speech segments of the investigator, pauses longer than 2 seconds, coughing, or any other irrelevant noise. The resulting sample is referred to as ‘the cleaned original sample’. Next, all stuttering-like dysfluencies are removed from the cleaned sample. Stuttering-like dysfluencies include part word repetitions, monosyllabic-word repetitions, prolongations, and blocks (the presence of dysfluencies, including phrase-initial blocks were discerned from the video recordings). The result of this procedure is a quasi-fluid speech sample. PDST is then calculated as:

\[
P DST = 100 \cdot \frac{\text{duration of the quasi-fluid speech sample}}{\text{duration of the cleaned original sample}}
\]

PDST scores for 3 participants revealed near zero levels in all conditions, reflecting no stuttering dysfluencies in their speech, and were excluded from statistical analysis.

---

3This duration threshold was based on findings that a silence interval (a pause) marking the intonation phrase boundary usually has a duration of 200 milliseconds, or more (Campione and Véronis, 2002; Fant et al., 2003).
6.3 Results

Group results

Statistical analysis (Friedman test) revealed a significant difference between the four experimental conditions for proportions of dysfluent phrase onsets (\( \chi^2 = 12.6; \) d.f. = 3; \( p = .006 \)). Conditions (b), (c), and (d) all reduced the proportions significantly in relation to the no-feedback condition (Wilcoxon signed ranks test, \( Z = 2.42; \) \( p = .015 \), \( Z = 2.549; \) \( p = .011 \), \( Z = 2.312; \) \( p = .021 \) respectively). There were further no significant differences between conditions (b), (c) and (d) \(^4\).

![Graph showing group means for proportions of dysfluent onsets across conditions (left), and for PDST measures across experimental conditions (right), z-scores normalized.](image)

Figure 6.2: Group means for proportions of dysfluent onsets across conditions (left), and for PDST measures across experimental conditions (right), z-scores normalized.

Group results for PDST measures show a similar pattern (Figure 6.2). Also with this variable, there was a significant effect of experimental condition (\( \chi^2 = 17.4; \) d.f. = 3; \( p = .001 \)), and a significant reduction of PDST scores in conditions (b),(c), and (d) in relation to the no-feedback condition (Wilcoxon signed ranks test, \( Z = 2.666; \) \( p = .008 \), \( Z = 2.429; \) \( p = .015 \), \( Z = 2.666; \) \( p = .008 \) respectively). There were no significant differences between conditions (b), (c) and (d). Comfortability scores revealed no clear preference between conditions (b), (c), and (d).

\(^4\)In order to display group means and individual trends in a meaningful way, the raw scores are normalized, so that each individual's trend across experimental conditions can be overlaid in the same numerical space. To achieve this, standard normal deviates (z-scores) were calculated for the raw scores for each individual speaker. The z transformation produces a distribution of scores with a mean of 0, and SD of 1, so that each score represents the distance in SD from the mean. After this transformation, a comparison can be drawn between the dysfluency trend of different speakers.
Individual results

Since group results in experiments with people who stutter show only a part of the picture, we present individual results for proportions of dysfluent onsets and for PDST measures (both as raw scores, as well as after z-scores normalization).

Although most participants show a reduction in conditions (b), (c), & (d) relative to the no-feedback condition (a), there are different trends of dysfluency across these conditions (Figure 6.3). For example, participants 8 and 12 have better reduced their dysfluencies on phrase onsets in condition (c) than in condition (b), and even more so in condition (d). The same time, participants 1 and 11 achieve a strong reduction in condition (c), but show more dysfluent onsets in condition (d). Participants 3 and 5 exhibit less reduction in condition (c) than in both conditions (b) & (d).

**Figure 6.3:** Individual results for proportions of dysfluent onset across conditions, presented as raw scores (left) & as z-scores normalized (right).

**Figure 6.4:** Individual PDST results across conditions, presented as raw scores (left) & as z-scores normalized (right).
The normalized individual PDST score in Figure 6.4 reveal that all participants improved their fluency level in conditions (b), (c), & (d). Two trends can be distinguished in the group. Participants 5, 7, and 10 have achieved more fluency in condition (c) than in condition (b), with even better or roughly the same level of fluency in condition (d). All other participants, however, show a pattern in which fluency levels are worse in condition (c) than in condition (b), while the level in condition (d) is either similar or better than in (c).

Sub-group results

An examination of the PDST scores under condition (a) in the left graph in Figure 6.4 reveals two rather distinct groups, with 5 participants having PDST scores lower than 10%, while the other 4 participants having PDST scores of 16% and higher. We term the first group as the ‘milder’ stuttering severity group, and the second as the ‘more severe’ group. Next, in order to compare the trends of both groups, we present z-scores normalized results for the two sub-groups.

![Figure 6.5: Sub-group results for proportions of dysfluent phrase onsets across conditions, z-scores normalized](image)

While in conditions (b) and (c) the more severe group seems to benefit from the feedback signal slightly more than the milder group, in condition (d) this trend is reversed (left graph in Figure 6.5). The milder group shows more reduction in dysfluent phrase onsets in condition (d) than the more severe group. A similar pattern can be seen in the sub-group results for PDST measures (right graph in Figure 6.5). Also here, the group of participants with a more severe dysfluency seem to improve their fluency levels to a greater extent than the other group only in conditions (b) and (c), but not in condition (d).
Participants' perceptions

We report a number of remarks expressed by participants in relation to the experimental conditions:

» Most participants agreed that condition (b), and DAF procedures in general, may be effective in reducing dysfluencies, but they do not help with the initiation of phrases, which remains an important concern.

» One participant remarked that the time when audio is fed back to the speaker in condition (c) can provide a ‘resting’ interval before initiating the next phrase, and therefore can help with maintaining fluency.

» Participants noted that it is difficult to predict the timing of feedback activation in condition (c), which may compromise its effectiveness to stabilize the speech production process.

» A number of participants noted that condition (d) is effective in helping to initiate phrases. Others, however, regarded the 1500 ms delay as being too slow, and not assisting with fluency since it is not possible to monitor oneself in real-time.

6.4 Discussion

Group trends

The group results for the proportion of dysfluent phrase onsets reveal that the adaptive feedback condition (c) significantly reduced the proportion of dysfluent onsets in relation to the baseline condition. However, the results did not support our hypothesis for a significant advantage of the adaptive feedback condition (c) over the standard DAF condition (b). Judging from Figure 6.2, there is a slight tendency for larger improvement in phrase initiation under condition (d). We interpret the finding that the adaptive feedback condition (c) did not provide significant advantage over the other experimental conditions in relation to the fact that the audio feedback in condition (c) was withdrawn once a phrase onset had been detected. If the audio cue was removed before speakers had the chance to produce an articulatory movement sufficient for resulting in a fluent syllable, the feedback procedure would not offer an efficient support for initiating phrases. If that is the case, then activating the standard DAF signal once phrase onset has been detected, as occurs in condition (d), would provide a continuation for the audio cue, and allow speakers to complete a fluent phrase initiation.

The group pattern for overall fluency results, measured by PDST scores, reveals a similar pattern, by which the adaptive feedback condition (c) significantly reduced PDST scores in relation to the baseline condition, but not significantly
better than other experimental conditions. In fact, Figure 6.2 reveals a tendency for conditions (b) and (d) to reduce PDST scores somewhat more than condition (c). These results seem to suggest that while speakers are able to utilize the DAF signal in mid-phrases, the additional audio cue during silences does not further enhance fluency. This ‘ceiling’ effect could be due to providing too much audio information during speech, when both mid-phrase and silences are filled with audio feedback. It seems that beyond some amount of auditory information, additional signals may be detrimental instead of beneficial. Alternatively, the finding that the additional audio cue during silences did not provide significant advantage may reflect the inadequate nature of the auditory cue used in this study.

Although comfortability scores revealed no clear preference between conditions (b), (c), and (d), participants tended to rate condition (b) as (slightly) more comfortable than condition (c). The reason could lie in the fact that most participants were familiar with a standard DAF procedure (either by experience or with the idea of it). As some participants have remarked, anticipating the nature of the audio signal in condition (b), could be making it more acceptable and less annoying than the less predictable audio signals in other conditions.

Individual variability

Individual data reveal high variability in participants’ responses to different feedback conditions, suggesting that different AAF procedures may have differential effects on various types of dysfluent speech. This idea is consistent with the notion that biological subtypes of stuttering may exist, and that different subtypes respond differently to AAF exposure. A similar idea was proposed in Foundas et al. (2004), who reported that DAF had a stronger fluency-inducing effect in the subgroup of stuttering individuals with rightward planum temporale asymmetry (i.e., right planum temporale larger than left planum temporale). In order to account for the high individual variability in response to AAF methods, it has been suggested that individual differences are influenced by the degree of auditory sensitivity (Howell et al., 2006), or by idiosyncratic abnormalities in the organization of auditory and motor related brain areas of people who stutter (Salmelin et al., 1998; Fox et al., 2000; Watkins et al., 2008).

Altogether, research findings point towards the existence of multi-causality in the etiology of stuttering, and therefore, of subtypes among individuals who stutter (Yairi, 2007).

An interesting trend was discovered when the group of participants was partitioned into sub-groups according to stuttering severity. Results for both the proportion of dysfluent phrase onsets and PDST scores, as can be seen in Figure 6.5, suggest that speakers with a higher level of dysfluency at baseline (no-feedback) condition have achieved larger relative dysfluency reductions compared to speakers with a milder PDST level of dysfluency at baseline. This pattern is consistent with earlier reports (Van Borsel et al., 2007). For example,
in Antipova et al. (2008), the authors reveal that the stuttering frequency of those subjects with more severe stuttering was reduced more than stuttering frequency of those with mild stuttering. The experience with our participants suggests that individuals with a milder dysfluency may not benefit as much from AAF procedures, since their speech motor system is relatively stable, and their level of confidence towards speech is relatively high. On the other hand, individuals with a more severe dysfluency, may be equipped with more fragile or sensitive speech motor systems (Van Lieshout et al., 2004), which could ‘gain’ more stability or reassurance from external cues.

Limitations of the current study

Individuals produce spontaneous speech in extremely variable manners, making experimental control more difficult for these tasks. It is, therefore, worthwhile to illustrate a number of potential sources of ambiguity related to the sampling of dysfluencies in spontaneous speech. First, speakers may be unconsciously using fluency techniques. Although we have explicitly instructed participants to not use any, at least two of them reported to have been using ‘slow speech’ and breathing techniques during the experimental tasks. Having internalized and automated fluency strategies to a certain degree, speakers may not be able to ‘switch them on’ upon command.

Second, when individuals who stutter produce a monologue on an engaging topic, they may place themselves in a ‘performativc’ mode, in which their speech pattern deviates from their habitual way of speech, leading to enhanced fluency. Two of our participants have associated ‘talking enthusiastically’ with enhanced fluency, while noting that this situation may not represent their ‘normal’ speech performance. This observation is in line with an account of fluency enhancement outlined in Alm (2005), which proposes that switching one’s speech manner from habitual to modified (as in imitation, acting, speaking with unnatural rhythm etc.) may enhance fluency through the de-automatization of speech motor control. According to this view, fluency is achieved by switching speech production from the supposedly dysfunctional medial system (the basal ganglia & supplementary motor area) responsible for automatized speech movements, to the lateral system (lateral premotor cortex & cerebellum) which coordinates de-automatized, and externally cued speech production (Alm, 2004).

Furthermore, dysfluencies seem to be strongly modulated by the presence of ‘di cult’ phonemes in the stream of speech, individual to each person. Anticipating those ‘fearful sounds’ is an integral part of the psychological experience of most people who stutter. Consequently, these speakers develop a variety of idiosyncratic strategies (which are distinct from ‘standard’ fluency techniques) for coping with the problematic sounds, such as reformulating phrases, substituting synonyms, and others. The influence of these ‘online speech manipulations’ on fluency may interact with the effects of external audio cues in an unpredictable way.
6.5 Conclusions and future directions

The adaptive feedback procedure used in this study resulted in strong reductions of dysfluencies at phrase onset positions. The results suggest that even higher reduction can be achieved if the audio feedback cue persists beyond the moment of phrase onset. Although the results indicate that an audio cue provided during silent segments may assist some speakers to initiate speech, they do not support the hypothesis that on a group level, this form of auditory cues result in a significantly larger fluency enhancement than the classic DAF procedure. The results for overall fluency levels reveal that an audio cue provided during silent segments leads to fluency improvement comparable with that of a standard DAF procedure. Investigation of individual results revealed a large degree of variability, confirming earlier evidence that persons who stutter react differently to AAF procedures.

At the outset of the study, it was assumed that, at least for some speakers who stutter, an adaptive presentation of feedback signals will appear as more comfortable than the continuously presented feedback signal in standard AAF procedures. However, considering the reactions of the study participants, it now seems that playing back a buffer of 1500 ms was not an optimal choice for an audio cue during silences. Hearing a segment this long seems to create a distracting effect in the process of speech production, and can be experienced as uncomfortable for some speakers. The distracting effect may be attributed to the interference with higher-level linguistic processes of utterance planning, which would normally take place during the silent pauses between phrases.

Within the framework of adaptive feedback procedures, audio cues can be considered which do not involve the straightforward playback of one's speech. Instead, an auditory cue can be designed which preserves the characteristics of the speaker's voice, but is devoid of linguistic content. Auditory cues can be further abstracted from the speaker's voice, and contain pre-recorded audio samples which can vary in the degree of linguistic content, from single syllables or single vowels, to non-linguistic cues, such as percussive sounds, or even synthetic tones and beats. There is, in fact, experimental evidence for fluency enhancing effects of non-linguistic auditory cues (Azrin et al., 1968; Martin and Haroldson, 1979; Howell and Archer, 1984), which could be used within an adaptive scheme. This approach would especially be relevant from a theoretical perspective which attributes fluency enhancing effects of auditory cues to influencing the timekeeping processes in speech production, thus not requiring linguistic content (Howell, 2004). However, following the insights from our study, it can be suggested that arriving at effective and comfortable feedback configurations can be best done on an individual basis.

Our experience with studying the effects of AAF procedures with people who stutter brings into question the group experimental design. The extent of individual variability (especially in spontaneous speech tasks) makes it difficult
to draw any solid conclusions on the group level. To be able to potentially observe group effects, a considerably large sample size would be needed, a demand which may slow down research activities. Another constraining factor is the standardization of AAF settings when studies focus on the group level. It has become a common understanding that individuals benefit most from customized AAF settings, and by standardizing these settings during studies, researchers may be overlooking the potential value of AAF procedures. Instead, research could focus on single-case studies where investigators strive to optimize AAF settings for each individual, and examine how performance with these settings changes over time and communicational contexts. This approach is echoed in the conclusions drawn by Lincoln et al. (2010), who propose that more useful information will be gained from carefully designed case studies that establish optimal AAF settings for individuals who stutter.

We believe that the current investigation provides a good starting point for further experimentations with variations of adaptive feedback procedures. There is little doubt that results can be improved, both in terms of fluency enhancement, and levels of comfortability with the procedure. Although we have outlined a number of potential caveats in conducting AAF experiments with spontaneous speech, the motivation for this approach must be put forward. The social validity of applying AAF procedures to realistic communication situations is of great importance (Lincoln et al., 2010), and demands a better understanding of fluency modulation in the context of spontaneous speech. Collecting participants’ impressions about their experience with the experimental conditions has been instrumental for a better understanding of the limitations and potentials of the proposed AAF procedures. Future attempts should take into account the aspect of individual variability, customization vs. standardization of AAF settings, and remain attentive to the experience of persons who stutter.

Despite the vast amount of research into stuttering during the last 40 years, there is still no consensus among researchers and clinicians (often decorated by heated debates) about the nature and etiology of the disorder, or the goals and methods of clinical intervention (Ratner, 2010). The widespread availability of aordable mobile technologies is likely to keep driving the development of a new generation of stuttering management devices. It is likely that future devices will be based on adaptive feedback procedures, with robust signal analysis capabilities, and elaborate customization options. In spite of the complexity involved, we strongly believe that it will be worthwhile to further explore adaptive feedback procedures, and devise methods for AAF procedures to be ‘tailored’ to individuals’ speakers needs.