Research Report

Emotional and physiological responses of fluent listeners while watching the speech of adults who stutter

Vijaya K. Guntupalli‡, D. Erik Everhart‡, Joseph Kalinowski§, Chayadevie Nanjundeswaran§ and Tim Saltuklaroglu†

‡Department of Communicative Disorders, East Tennessee State University, Johnson City, TN, USA
§Department of Psychology, East Carolina University, Greenville, NC, USA
Department of Communication Sciences and Disorders, East Carolina University, Greenville, NC, USA
¶Department of Audiology and Speech Pathology, University of Tennessee, Knoxville, TN, USA

(Received 1 June 2005; accepted 2 February 2006)

Abstract

Background: People who stutter produce speech that is characterized by intermittent, involuntary part-word repetitions and prolongations. In addition to these signature acoustic manifestations, those who stutter often display repetitive and fixated behaviours outside the speech producing mechanism (e.g. in the head, arm, fingers, nares, etc.). Previous research has examined the attitudes and perceptions of those who stutter and people who frequently interact with them (e.g. relatives, parents, employers). Results have shown an unequivocal, powerful and robust negative stereotype despite a lack of defined differences in personality structure between people who stutter and normally fluent individuals. However, physiological investigations of listener responses during moments of stuttering are limited. There is a need for data that simultaneously examine physiological responses (e.g. heart rate and galvanic skin conductance) and subjective behavioural responses to stuttering. The pairing of these objective and subjective data may provide information that casts light on the genesis of negative stereotypes associated with stuttering, the development of compensatory mechanisms in those who stutter, and the true impact of stuttering on senders and receivers alike.

Address correspondence to: Vijaya Guntupalli, Department of Communicative Disorders, East Tennessee State University, Lamb Hall, Box 70643, Johnson City, TN 37614, USA; e-mail: vkg0404@ecu.edu
Aims: To compare the emotional and physiological responses of fluent speakers while listening and observing fluent and severe stuttered speech samples.

Methods & Procedures: Twenty adult participants (mean age = 24.15 years, standard deviation = 3.40) observed speech samples of two fluent speakers and two speakers who stutter reading aloud. Participants' skin conductance and heart rate changes were measured as physiological responses to stuttered or fluent speech samples. Participants' subjective responses on arousal (excited–calm) and valence (happy–unhappy) dimensions were assessed via the Self-Assessment Manikin (SAM) rating scale with an additional questionnaire comprised of a set of nine bipolar adjectives.

Outcomes & Results: Results showed significantly increased skin conductance and lower mean heart rate during the presentation of stuttered speech relative to the presentation of fluent speech samples ($p < 0.05$). Listeners also self-rated themselves as being more aroused, unhappy, nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassed, and annoyed while viewing stuttered speech relative to the fluent speech.

Conclusions: These data support the notion that stutter-filled speech can elicit physiological and emotional responses in listeners. Clinicians who treat stuttering should be aware that listeners show involuntary physiological responses to moderate–severe stuttering that probably remain salient over time and contribute to the evolution of negative stereotypes of people who stutter. With this in mind, it is hoped that clinicians can work with people who stutter to develop appropriate coping strategies. The role of amygdala and mirror neural mechanism in physiological and subjective responses to stuttering is discussed.

Keywords: stuttering, emotions, attitudes, skin conductance, heart rate, stuttering therapy, mirror neurons.

What this paper adds

Previous research has shown that people who stutter were perceived as being nervous, tense, shy, quiet, reticent, guarded, avoiding, introverted, afraid, passive, self-derogatory, and more sensitive relative to people who do not stutter. The present study found that fluent listeners rated themselves as more unhappy, aroused, nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassed, and annoyed while viewing and listening to stuttered speech relative to fluent speech, and showed increased skin conductance and reduced heart rate to stuttered speech as compared with fluent speech.

Introduction

It is likely that if you have never been regarded as a stutterer, you can come nowhere near appreciating the uncanny, crushing power of the social disapproval of whatever is regarded as stuttering. It is probably one of the most frightening, perplexing, and demoralizing influences to be found in our culture.

(Johnson 1946: 458)

Stuttered speech is filled with intermittent sound prolongations and/or rapid part-word repetitions. Oftentimes, these aberrant acoustic behaviours are associated with
intermittent physical tension and struggle behaviours such as head jerks, arm jerks, finger tapping, excessive eye-blinks, wrinkling of the forehead, lip bites, clavicular breathing, and involuntary arm, torso, and leg movements. The use of these struggle behaviours outside the vocal tract is suggestive of primitive evolutionary survival responses in oral and manual mechanisms (Ferrari et al. 2003) that are still housed deep within the neural systems of all who speak. The inability to express thoughts fluently oftentimes brings shame, guilt, and embarrassment to the individuals who stutter influencing their social and vocational choices (Van Riper 1971, Ginsberg 2000, Kalinowski and Saltuklaroglu 2005). Additionally, the jarring nature of acoustic and physical manifestations that often accompanies moderate–severe stuttering may induce negative emotional responses in listeners, which alters communication between the person who stutters and their listeners (Bloodstein 1995, Guntupalli et al. 2006). Persons who stutter often report negative emotional responses (e.g. impatience, giggling, embarrassment, surprise, pity or laughter) from listeners, especially, if they are strangers. These negative emotional responses explain in part the development of compensatory strategies such as avoidance of sounds, words, people and places.

Successful communication may be described as a free-flowing exchange of linguistic and emotional information between individuals as they switch effortlessly speaker–listener roles. However, when stuttering manifests itself, an unnatural break is introduced into the natural ebb and flow of information exchange. The unprepared listener might show a natural startle reaction in response to the sudden aberrant stuttering behaviours. When listener observes an unexpected behaviour (such as stuttering), there may be a surprise or an illusionary threat response to the stimulus, but when this illusion disappears with the realization of no imminent threat, there is an emotional incongruity, leading to a giggle or laugh from the listener (Ramachandran 1998). Rosenberg and Curtiss (1954) observed that listeners became much less mobile, lost eye contact, and reduced their speech output when confronted with a person who stutters. McDonald and Frick (1954) examined the attitudes of store clerks while speaking to a person who stutters using a 25-item questionnaire with yes/no type answers. The findings of the study revealed that listeners expressed sympathy, embarrassment, curiosity and surprise when talking to a person who stutters. Together these findings suggest that listeners generally appear to possess unfavourable perceptions towards individuals who stutter.

Woods and Williams (1976) investigated how various groups rated hypothetical individuals who stutter on a semantic differential scale compared to hypothetical individuals who did not stutter. Results demonstrated that individuals who stutter were perceived to be nervous, tense, shy, quiet, reticent, guarded, avoiding, introverted, afraid, passive, self-derogatory and sensitive. These negative attributes given to hypothetical individuals who stutter were robust and found in many groups including students (St Louis and Lass 1981, White and Collins 1984, Dorsey and Guenther 2000), parents (Woods and Williams 1976, Crowe and Cooper 1977, Fowlie and Cooper 1978), speech–language clinicians (Woods and Williams 1971, Yairi and Williams 1970, Turnbaugh et al. 1979, Cooper and Cooper 1985, Cooper and Rustin 1985, Lass et al. 1989, Kalinowski et al. 1993, Rami et al. 2003), teachers and professors (Crowe and Walton 1981, Yeakle and Cooper 1986, Lass et al. 1992, Dorsey and Guenther 2000), special educators (Ruscello and Lass 1994), people who stutter (Kalinowski et al. 1987, Lass et al. 1995) and vocational rehabilitation counsellors (Hurst and Cooper 1983).
Given these findings, ample evidence exists to indicate that many listeners have developed definitive opinions about the personality characteristics of people who stutter in spite of lack of empirical evidence suggesting a personality difference that is typical of people who stutter as a group (Bloodstein 1995). It is suspected that listeners developed these negative stereotypical attitudes from their affective or emotional responses while witnessing moments of stuttering (Kalinowski et al. 1987, Guntupalli et al. 2006). Listeners’ affective feelings are thought to play an important role in the early conception of social groups. In other words, when a particular group of people or a stimulus possesses the capability of evoking emotional responses in the listener, these feelings may transfer into attitudes about the group (Bennett and Hacker 2003). Simply put, if stuttered speech behaviours induces an emotional response in listeners, then those feelings may be transferred to the group of people (i.e. people who stutter) producing them.

Empirical evidence exists to suggest that 'emotional response' to pictures or film clips are fundamentally organized according to two primitive motivational factors: affective dimensions of valence (pleasant–unpleasant) and arousal (excited–calm) (Lang et al. 1993, Bradley et al. 2001). That is, emotional responses such as disgust, sadness, fear etc. are determined by affective evaluation of the stimuli on arousal and valence dimensions. It has also been found that several physiological responses such as skin conductance, heart rate, facial electromyography and startle reflex eye-blink magnitude covary with the affective dimension of valence and arousal. Among these physiological measures, skin conductance response (SCR) is considered to be a sensitive and reliable measure of autonomic arousal to various emotional or affective stimuli (Lang et al. 1993, Ashcroft et al. 2001, Dietrich and Roaman 2001). A number of studies have reported increased skin conductance with self-reports of arousal to emotional stimuli, independent of whether the emotional experience is reported to be pleasant or unpleasant, indicating that SCR is a good predictor of autonomic arousal to emotional stimuli (Lang et al. 1993, Palomba et al. 2000, Bradley et al. 2001, Frazier et al. 2004, Gomez et al. 2005). Heart rate (HR) is another physiological response that is commonly associated with the valence of film stimuli. In general, affective valence plays a greater role in determining cardiac responses and mean heart rate is reported to be lower (deceleration) for unpleasant film stimuli relative to pleasant film stimuli (Lang et al. 1993, Palomba et al. 1997, Gomez et al. 2005). However, greater variability in HR pattern (e.g. acceleration or deceleration) is noted depending on the type of emotional stimuli (Stern et al. 2001). Palomba et al. (2000) reported an increase in SCR and cardiac deceleration while watching a surgery film whereas an increase in SCR and cardiac acceleration was noted during a threat film suggesting that autonomic response patterns are related to the specific content of the unpleasant films (i.e. threat versus surgery).

To the best of our knowledge, Guntupalli et al. (in press) was the only study that investigated the psychophysiological responses of college students while they watched and listened to individuals who stutter and normally fluent individuals reading aloud. Results revealed that listeners showed increased skin conductance and reduced heart rate while watching the stuttered speech samples relative to the fluent speech samples suggesting that listeners were physiologically aroused and support attentive in response to. These preliminary findings using physiological measures provide evidence to stuttered stimuli. The notion that aberrant stuttering behaviours have an impact on listeners at a physiological level. The authors further hypothesized that aberrant and anomalous stuttering behaviours probably stimulate
the ‘mirror neuronal mechanism’ in listeners eliciting an experiential sense of aberrant stuttering moments. ‘Mirror neurons’ are thought to be the fundamental mechanism for facilitating the understanding of the actions and emotions of others by internally simulating or mimicking those actions. This internal mimicry provides neural stimulation of an action or emotion. Thus, during observation of an action or emotional state, a direct neural link is created between sender and receiver (Rizzolatti et al. 1996, Carr et al. 2003, Gallese et al. 2004, Leslie et al. 2004). According to this hypothesis, observers are not only seeing or hearing actions, but may also experience the emotional state associated with those actions as if they were performing similar actions (Gallese et al. 2004).

The preliminary findings using physiological measures provide evidence to the notion that aberrant stuttering behaviours have an impact on listeners at a physiological level. This involuntary physiological arousal may suggest the uneasiness and discomfort that listeners might have towards the unexpected stuttering behaviour. To understand further the listener’s emotional reactions and to provide evidence to the observed physiological response, real-time evaluations of both emotional responses using self-report measures and physiological responses is necessary. Thus, the purpose of the present study was (1) to evaluate the listeners’ self-rated subjective valence and arousal ratings to stuttered and fluent speech, (2) to measure changes in skin conductance and heart rate of listeners’ while watching stuttered and fluent speech samples, and (3) to evaluate the participant’s subjective feelings on nine bipolar adjectives to fluent and stuttered speech samples.

**Methods**

**Participants**

Twenty normally fluent adults (ten males and ten females), aged 18–32 years (mean = 24.15, standard deviation (SD) = 3.40) participated in the study. Among the participants, 17 were students, one was in the armed forces, and the remaining two were managers employed full-time. All the participants reported having normal hearing and normal or corrected vision. They all met the following selection criteria: (1) no training in the area of speech, language and hearing disorders, (2) no self-report of a family member or acquaintance with communication disorder(s), including fluency disorders, (3) no self-report of speech, language, or hearing impairment(s), and (4) they were all native speakers of English.

**Stimuli**

The stimuli consisted of four video clips with two fluent and two stuttered speech samples of 30 s. The original speech samples were recorded in a sound attenuated room using a digital video camera (JVC miniDV GR-D70U) from two individuals who stutter and two individuals who do not stutter while they read a junior high school level passage. All four individuals were white males and aged between 18 and 24 years. From these speech samples, 30 s of unedited fluent and stuttered speech tokens were copied onto a DVD for playback to the listeners. Two certified speech language pathologists who have experience in stuttering therapy rated the severity of stuttering. Both speech samples with stuttering were rated as ‘severe’ (i.e. sample 1
with a score of 36, and sample 2 with a score of 32 using SSI-3; Riley 1994). Consistent with severity ratings, both individuals exhibited high levels of stuttering moments filled with repetitions and prolongations of speech sounds along with the postural fixations. Specifically, the average frequency of stuttered syllables for samples one and two were 45 and 21%, respectively. For the frequency measure, part-word repetitions, prolongations and postural fixations on speech sounds were counted. In addition, the first stuttered sample was characterized by excessive head jerking and hand movements near the mouth whereas the second stuttered sample was characterized by head jerking and eye closure.

**Apparatus**

The speech samples were played through a Panasonic DVD player (model number DMR-E30). The video output from the DVD player was routed through a four-way audio/video selector (Radio shack, model number 05A04), before presenting to the participants via a Glasstron Home Theatre System (SONY personal LCD monitor PLM-A35), consisting of a personal monitor with special viewing lenses for video presentation (Saltuklaroglu et al. 2004). Video output from the DVD player was also sent to the colour television for the experimenter to view and monitor the stimulus. The audio output from the DVD player was routed through a Mackie 1202 mixer and presented to the listeners via E-A-R 3A insert earphones at a comfortable listening level. PSYLAB computer instrumentation (Contact Precision Instruments, Cambridge, MA, USA) was used to record the autonomic responses (i.e. SCR and HR). A personal laptop computer (Compaq Presario 2100) was used to present the rating scales electronically. The video output from the laptop computer was routed through the same four-way audio/video selector and finally to the participants via Glasstron home theatre system.

**Response measurement**

**Self-report measures**

Affective valence and arousal ratings were obtained using the paper-and-pencil version of Lang’s Self-Assessment Manikin (SAM; Bradley and Lang 1994). Valence was rated by marking on five graphics depicting the manikin with facial expressions ranging from a broad smile to a severe frown. The valence dimension extends from pleasant, happy, or positive valence at one end to unpleasant, unhappy, unsatisfied or negative valence at the other. Arousal was rated similarly using five graphics depicting the manikin at different levels of visceral agitation ranging from completely stimulated, excited, or aroused to completely relaxed, calm, or unaroused. The participants could select any of the five figures comprising each scale, which results in a five-point rating scale for each dimension. Ratings were scored such that 5 represents a high rating on each dimension (i.e. high pleasure, high arousal), and 1 represents a low rating on each dimension (i.e. low pleasure, low arousal). These two dimensions, valence and arousal, account for most of the independent variance in emotional responses (Greenwald et al. 1989, Lang et al. 1993). In addition to the valence and arousal ratings to the stimuli, a participant’s self-rated affective feelings were obtained using nine bipolar adjectives on a nine-point Likert scale (see
appendix A for the rating scale). Nine bipolar adjectives were appropriately selected to evaluate the subjective feelings of listeners.

Physiological measures

Skin conductance responses to speech stimuli were recorded using a pair of 8-mm$^2$ pre-wired Ag/AgCl electrodes (TD-22 EL1, Med Associates Inc. St Albans, VT, USA) attached to the palmer surface of the distal phalanges of the second and third fingers of the left hand. The PSYLAB SC5 skin conductance coupler provides a 24-bit accuracy A–D converter built into the Amplifier. SC5 coupler provides a 0.5-V constant voltage across the two electrodes and measures the skin conductance directly. The accuracy of skin conductance measurement was checked with the two calibrator buttons (20 and 0.1 $\mu$S) provided in the pre-amplifier.

Heart rate or inter-beat interval measures were obtained using electrocardiogram (ECG). ECG was recorded using two disposable Ag/AgCl electrodes (3M Red dot monitoring electrodes with solid gel, 2239) placed one at the base of neck on the right side and other on the left ankle. ECG was recorded with Bioamplifier with the high pass filter set at 10 Hz and the low pass filter set at 40 Hz. Gain was adjusted to 1 mV with the Hum filter on. The output from the bioamplifier was fed to the interval timer to measure the inter-beat interval (time between two successive R-waves). The inter-beat interval was measured using a single millisecond accuracy.

Procedure

Participants were provided with a general description of the stimuli, the rating task, and the recording techniques and they were asked to sign an informed consent form approved by the UMCIRB, East Carolina University. Two investigators were present in the room throughout the experiment. At the beginning of the experiment, participants were asked to wash their hands with soap and water and dry them thoroughly. The purpose of this was to ensure an equal degree of skin hydration across all participants for the measurement of skin conductance. The participants were seated in a comfortable chair facing away from the equipment. They were then given a baseline-rating task. During this, participants rated their responses to a hypothetical person who stutters and a hypothetical person who does not stutter on the SAM valence and arousal scale and on the nine bipolar adjectives. Order of completion of SAM scales and questionnaire was counterbalanced.

Following the completion of the baseline-rating task, the electrodes for collecting autonomic responses data (i.e. SCR and inter-beat interval) were applied to the appropriate areas. To view and hear the speech tokens, the Sony Glasstron Home Theatre System was placed over the participants’ eyes and E-A-R insert earphones were placed in the ears of the participant using disposable foam ear tips. All of the participants were given a 5-min adaptation period before viewing the first experimental video stimulus. They were asked to remain in a comfortable and relaxed state. Following the adaptation period, the investigator collected baseline autonomic responses for 30 s while the participant remained in a relaxed state before being presented with the first speech token. SCR and inter-beat intervals during the baseline and experimental tasks were continuously recorded throughout the
procedure. However, the investigator placed event markers in order to indicate the onset of stimulus presentation.

After viewing each speech token, participants were asked to provide their affective responses on the SAM scale of valence and arousal dimensions and to nine bipolar adjectives. The scales were presented electronically as PowerPoint slides (via the Glasstron Home Theater System) and participants were asked to provide their responses orally and the investigator recorded these responses on paper. This was done to minimize the movement artefacts associated with writing, removing and placing the Glasstron Home Theater System. Following the rating task to speech token, a 2-min recovery period was given before measuring the baseline for the next speech token. The 2-min period was given to allow physiological measures to return to baseline levels. The fluent and stuttered speech tokens were arranged using diagram-balanced Latin square method and participants were randomly assigned to one order.

Analysis was completed for a 60-s window that consisted of 30-s period before the presentation of the stimulus (baseline) and 30-s period during the presentation of the stimulus (response). The average amplitude of skin conductance (in μS) and average inter-beat interval (in ms) during the baseline and also while viewing the speech stimulus was measured for all the stimulus tokens. Further, the differences between baseline and stimulus response was calculated for skin conductance and heart rate changes for all speech samples. Descriptive and inferential statistical analyses were performed using SPSS software (PC version 11.0).

Results

Self-report measures

The mean and standard errors of ratings on affective valence and arousal dimensions and other nine bipolar adjectives to fluent and stuttered speech samples are shown in table 1. Here, N0, N1, N2 denote a participant’s self-ratings at the baseline, first fluent and second fluent speech samples, respectively. Similarly, S0, S1, S2 denote a participant’s self-ratings at baseline, first stuttered and second stuttered samples, respectively.

A 2 × 3 repeated-measures analysis of variance (ANOVA) was performed for each rating scale to investigate the differences in rating scores as a function of group (fluent versus stuttered) and order of stimulus (i.e. baseline versus first versus second speech sample). Eleven 2 × 3 (group × order) repeated-measure ANOVA's for the 11 rating scales revealed a significant main effect of group (fluent versus stuttered speech) for all the ratings scales (p<0.05). Except for two rating scales (funny–sad and emotional–bland), there was a significant order effect (baseline, first and second speech samples) for the remaining nine rating scales (p<0.05). However, there was no significant interaction between groups by order (p>0.05) (table 2).

Five planned single-degrees of freedom (d.f.) contrasts were performed for each rating scale to evaluate the differences in means ratings between (1) baseline (N0) versus combined ratings of N1 and N2, (2) baseline (S0) versus combined ratings of S1 and S2, (3) N1 versus N2, (4) S1 versus S2, and (5) combined ratings of N1 and N2 versus combined ratings of S1 and S2. For all rating scales, a significant difference was found between combined ratings of N1 and N2 versus combined ratings of S1 and S2 (p<0.05). Additionally, a significant difference was observed
Table 1. Mean and standard error of ratings ($n=20$) to fluent and stuttered speech samples as a function of 11 rating scales

<table>
<thead>
<tr>
<th>Rating scale</th>
<th>N0 (SD)</th>
<th>N1 (SD)</th>
<th>N2 (SD)</th>
<th>S0 (SD)</th>
<th>S1 (SD)</th>
<th>S2 (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence (happy–unhappy)</td>
<td>1.55 (0.17)</td>
<td>1.70 (0.16)</td>
<td>1.85 (0.20)</td>
<td>2.65 (0.18)</td>
<td>3.20 (0.22)</td>
<td>3.45 (0.17)</td>
</tr>
<tr>
<td>Arousal (excited–calm)</td>
<td>4.55 (0.14)</td>
<td>4.25 (0.16)</td>
<td>4.25 (0.20)</td>
<td>3.40 (0.15)</td>
<td>3.00 (0.21)</td>
<td>3.15 (0.24)</td>
</tr>
<tr>
<td>Nervous (1)–calm (9)</td>
<td>8.15 (0.23)</td>
<td>7.55 (0.27)</td>
<td>7.65 (0.26)</td>
<td>5.65 (0.46)</td>
<td>4.80 (0.43)</td>
<td>5.20 (0.47)</td>
</tr>
<tr>
<td>Comfortable (1)–uncomfortable (9)</td>
<td>1.85 (0.29)</td>
<td>2.65 (0.37)</td>
<td>2.55 (0.38)</td>
<td>4.65 (0.42)</td>
<td>4.95 (0.44)</td>
<td>5.25 (0.49)</td>
</tr>
<tr>
<td>Funny (1)–sad (9)</td>
<td>4.05 (0.29)</td>
<td>4.75 (0.22)</td>
<td>4.55 (0.20)</td>
<td>5.55 (0.34)</td>
<td>6.05 (0.33)</td>
<td>5.90 (0.32)</td>
</tr>
<tr>
<td>Tense (1)–relaxed (9)</td>
<td>8.10 (0.28)</td>
<td>7.25 (0.31)</td>
<td>7.45 (0.32)</td>
<td>5.05 (0.41)</td>
<td>4.60 (0.47)</td>
<td>4.95 (0.45)</td>
</tr>
<tr>
<td>Pleasant (1)–unpleasant (9)</td>
<td>2.35 (0.41)</td>
<td>3.10 (0.41)</td>
<td>2.90 (0.49)</td>
<td>4.50 (0.36)</td>
<td>5.45 (0.44)</td>
<td>5.95 (0.32)</td>
</tr>
<tr>
<td>Avoiding (1)–approaching (9)</td>
<td>7.45 (0.32)</td>
<td>6.95 (0.34)</td>
<td>7.25 (0.32)</td>
<td>5.55 (0.37)</td>
<td>4.55 (0.32)</td>
<td>4.45 (0.41)</td>
</tr>
<tr>
<td>Emotional (1)–bland (9)</td>
<td>5.75 (0.44)</td>
<td>6.40 (0.46)</td>
<td>5.95 (0.36)</td>
<td>4.10 (0.32)</td>
<td>3.60 (0.33)</td>
<td>3.65 (0.31)</td>
</tr>
<tr>
<td>Embarrassing (1)–soothing (9)</td>
<td>6.95 (0.34)</td>
<td>6.60 (0.29)</td>
<td>6.70 (0.35)</td>
<td>5.35 (0.34)</td>
<td>4.60 (0.26)</td>
<td>4.30 (0.33)</td>
</tr>
<tr>
<td>Annoyed (1)–joyful (9)</td>
<td>7.15 (0.33)</td>
<td>5.90 (0.32)</td>
<td>5.90 (0.34)</td>
<td>5.45 (0.30)</td>
<td>4.50 (0.25)</td>
<td>4.50 (0.32)</td>
</tr>
</tbody>
</table>

N0, N1 and N2 denote responses at the baseline, first and second fluent speech sample samples, respectively. S0, S1 and S2 denote responses at the baseline, first and second stuttered speech samples, respectively.
between baseline ratings (S0) versus combined ratings of S1 and S2 for ratings on valence, arousal, pleasant–unpleasant, avoiding–approaching, embarrassing–soothing, and annoyed–joyful dimensions. Finally, a significant difference was found between the baseline ratings (N0) versus combined ratings on N1 and N2 ($p<0.05$) for ratings on arousal, nervous–calm, tense–relaxed, and annoyed–joyful dimensions.

### Physiological measures

The mean and standard errors (SEs) for the change in skin conductance values (response – baseline) while viewing the first and second fluent and first and second stuttered speech samples were 0.378 µS (SE=0.07), 0.137 µS (SE=0.07), 0.569

<table>
<thead>
<tr>
<th>Rating scale</th>
<th>d.f.</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
<th>$\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence ratings: Group</td>
<td>1, 19</td>
<td>67.57</td>
<td>&lt;0.001*</td>
<td>0.78</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>10.58</td>
<td>&lt;0.001*</td>
<td>0.36</td>
<td>0.98</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>2.36</td>
<td>0.109</td>
<td>0.11</td>
<td>0.44</td>
</tr>
<tr>
<td>Arousal ratings: Group</td>
<td>1, 19</td>
<td>38.67</td>
<td>&lt;0.001*</td>
<td>0.67</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>5.38</td>
<td>&lt;0.009*</td>
<td>0.22</td>
<td>0.81</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.09</td>
<td>0.889</td>
<td>0.01</td>
<td>0.06</td>
</tr>
<tr>
<td>Nervous–calm: Group</td>
<td>1, 19</td>
<td>50.27</td>
<td>&lt;0.001*</td>
<td>0.73</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>6.35</td>
<td>0.004*</td>
<td>0.25</td>
<td>0.88</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.24</td>
<td>0.755</td>
<td>0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>Comfortable–uncomfortable: Group</td>
<td>1, 19</td>
<td>39.37</td>
<td>&lt;0.001*</td>
<td>0.66</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>4.24</td>
<td>0.022*</td>
<td>0.18</td>
<td>0.71</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.45</td>
<td>0.632</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>Funny–sad: Group</td>
<td>1, 19</td>
<td>14.09</td>
<td>0.001*</td>
<td>0.43</td>
<td>0.95</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>3.14</td>
<td>0.055</td>
<td>0.14</td>
<td>0.57</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.14</td>
<td>0.87</td>
<td>0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Tense–relaxed: Group</td>
<td>1, 19</td>
<td>50.18</td>
<td>&lt;0.001*</td>
<td>0.73</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>3.96</td>
<td>0.037*</td>
<td>0.17</td>
<td>0.61</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.64</td>
<td>0.536</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>Pleasant–unpleasant: Group</td>
<td>1, 19</td>
<td>38.21</td>
<td>&lt;0.001*</td>
<td>0.67</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>4.72</td>
<td>0.026*</td>
<td>0.20</td>
<td>0.66</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>1.56</td>
<td>0.224</td>
<td>0.08</td>
<td>0.31</td>
</tr>
<tr>
<td>Avoiding–approaching: Group</td>
<td>1, 19</td>
<td>46.74</td>
<td>&lt;0.001*</td>
<td>0.71</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>10.33</td>
<td>&lt;0.001*</td>
<td>0.35</td>
<td>0.98</td>
</tr>
<tr>
<td>Group x Order</td>
<td>2, 38</td>
<td>1.55</td>
<td>0.224</td>
<td>0.08</td>
<td>0.31</td>
</tr>
<tr>
<td>Emotional–bland: Group</td>
<td>1, 19</td>
<td>29.97</td>
<td>&lt;0.001*</td>
<td>0.61</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>0.43</td>
<td>0.654</td>
<td>0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>2.54</td>
<td>0.092</td>
<td>0.12</td>
<td>0.48</td>
</tr>
<tr>
<td>Embarrassing–soothing: Group</td>
<td>1, 19</td>
<td>40.00</td>
<td>&lt;0.001*</td>
<td>0.68</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>4.18</td>
<td>0.023*</td>
<td>0.18</td>
<td>0.70</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>1.61</td>
<td>0.214</td>
<td>0.08</td>
<td>0.32</td>
</tr>
<tr>
<td>Annoyed–joyful: Group</td>
<td>1, 19</td>
<td>36.13</td>
<td>&lt;0.001*</td>
<td>0.66</td>
<td>1.00</td>
</tr>
<tr>
<td>Order</td>
<td>2, 38</td>
<td>7.96</td>
<td>0.002*</td>
<td>0.30</td>
<td>0.93</td>
</tr>
<tr>
<td>Group x order</td>
<td>2, 38</td>
<td>0.32</td>
<td>0.727</td>
<td>0.02</td>
<td>0.10</td>
</tr>
</tbody>
</table>

*Significant at $p<0.05$; $p$-values are reported following a Huynh–Felt correction; effect size is indexed by $\eta^2$; power is indexed by $\Phi$ at $\alpha=0.05$. 

Vijaya K. Guntupalli et al.
For pooled data, a paired sample t-test (two-tailed) between the fluent and stuttered speech samples revealed a significant difference for change in skin conductance values [t(39)=2.34, p=0.024, \( \eta^2=0.19, \phi=0.52 \) at \( \alpha=0.05 \)].

The inter-beat interval was reduced to heart rate (beats per minute, bpm) changes for convenience and easy understanding. Heart rate was calculated for each condition by dividing 60 000 (i.e. the number of milliseconds per 1 min) by the inter-beat interval. The mean and standard errors for the change in heart rate (response – baseline) while viewing the first and second fluent and first and second stuttered speech samples were 1.70 bpm (SE=0.82), 2.74 bpm (SE=0.72), 0.48 bpm (SE=0.76), and −0.24 bpm (SE=0.59), respectively. For pooled data, a paired sample t-test (two-tailed) between fluent and stuttered speech samples revealed a significant difference for change in heart rate [t(33)=3.38, p=0.002, \( \eta^2=0.33, \phi=0.89 \) at \( \alpha=0.05 \)].

**Discussion**

The important findings of the present study are as follows. First, the stuttered speech samples elicited significantly increased skin conductance and lower mean heart rate relative to the fluent samples. Second, fluent participants self-perceived themselves as being more unhappy, aroused, nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassed, and annoyed while viewing and listening to stuttered speech relative to fluent speech. Third, participants had negative emotional responses even when they are asked to imagine watching a hypothetical individual who stutters (i.e. during baseline condition). Finally, participants' subjective responses did not differ significantly during the presentation of second stuttered and fluent speech sample compared to the first stuttered and fluent speech sample, respectively.

The finding that stuttering behaviours elicit an increased skin conductance and reduced heart rate relative to fluent speech in listeners who report no personal history or family history of communication disorders is consistent with the findings of a previous study (Guntupalli et al. in press), where they have used a similar experimental protocol. These findings indicate that when listeners are exposed to moderate–severe moments of stuttering, the perception of abrupt oscillatory or fixated speech disruptions can rapidly impact the autonomic nervous system causing an immediate autonomic arousal. It is hypothesized that the amygdala plays a major role in the initial surprise or fear that listeners experience as a result of the ‘sudden rush’ from the sensory aberrance.

The stimuli used in this study were characterized by moments of severe stuttering for the purpose of determining their effects on specific physiological measures (e.g. SC, and HR) in non-stuttering participants. However, even in the most severe cases of stuttering, overt speech disruptions are interspersed with periods of fluency. That is, unlike most communicative disorders, stuttering presents itself intermittently, with involuntary speech disruptions manifesting rapidly and abruptly among fluent utterances. The abrupt and disruptive nature of stuttering on communication is acute, drastic and powerful, creating an assault to the senses of unsuspecting naïve listeners. This unexpected sensory assault immediately leads to autonomic reactions in listeners that are possibly mediated by the activation of
amygdala, as evidenced by changes in skin conductance (SC) and HR. Listeners’ immediate reaction to the unexpected stuttering behaviour may be considered analogous to surprising a person by jumping out of a closet. Due to the primitive nature of repetitive gestural sounds seen in stuttering, the listener may perceive this as a simple startle or threat. In the evolutionary scheme, both sender and receiver have inherited the capacity to judge the intent of one another’s actions and behaviours. When the receiver is confronted with a person who stutters (i.e. sender), the aberrant and powerful actions of stuttering induces a surprise or threat causing the activation of autonomic nervous system. When listeners come to realize that the act of stuttering, although initially disturbing, is now non-threatening, an immediate relief from threat may bring about a giggle or laugh in the listener (Ramachandran 1998). Alternatively, the findings of increased skin conductance and reduced heart rate to stuttering behaviours could be because of attention being paid to the aberrant stuttering moments. The differential response (increase in skin conductance and decrease in heart rate) can be attributed to the orienting behaviour or attention paid to the stimulus by listeners (Stern et al. 2001).

Participants’ subjective responses to arousal and valence dimensions of SAM indicated that they self-perceived themselves as being significantly aroused/excited and unpleasant/unhappy in response to witnessing stuttering behaviours. The observed affective responses to stuttered speech samples relative to fluent speech samples on the SAM scales are in agreement with the observed physiological responses (i.e. increase in skin conductance and lower heart rate) to aberrant stuttering moments. In general, increased skin conductance is associated with arousing stimuli compared to calming stimuli and reduced heart rate is often associated with negative valence relative to positive valence (Lang et al. 1993, Palomba et al. 2000, Bradley et al. 2001, Gomez et al. 2005). Additionally, on the nine bipolar adjectives, participants self-rated themselves being more nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassed, and annoyed while viewing and listening to stuttered speech samples relative to fluent samples. The present findings are the first to suggest listeners have a negative emotional response to stuttering behaviours.

We hypothesize that when listeners’ passively view stuttered speech, these events can be mirrored in the listener’s brain giving them an experiential sense of stuttering behaviours and associated emotional response. Neurophysiological evidence has shown selective neural responses in superior temporal sulcus (STS) to the movement of faces, and even processes details such as shape of the mouth and direction of gaze (Perrett et al. 1984, Allison et al. 2000, Grossman et al. 2000) suggesting that the STS may decode the complex oro-facial movements and controls the affective responses through its outputs to the limbic system, frontal and parietal systems (Puce and Perrett 2004). Additionally, using functional magnetic resonance image (fMRI), researchers have observed that ‘observation’ as well as ‘imitation’ of emotional facial expressions of others elicits activity in the same cortical circuit. This circuit consists of the inferior frontal lobe, posterior parietal cortex and superior temporal sulcus, which are thought to be are important for action representation. These regions, along with the insula and amygdala are also thought to facilitate the understanding of feelings of others (i.e. empathy) or emotional experience associated with specific movements (Carr et al. 2003, Wicker et al. 2003). The stuttering moment creates an action or emotional state with a direct neural link between sender and receiver (Rizzolatti et al. 1996, Carr et al. 2003, Gallese et al. 2005).
Receivers are not only observing and listening to actions, but may also experience the emotional state associated with those actions as if they were performing similar actions (Gallese et al. 2004). These results showed similar negative emotional responses even when participants were asked to imagine watching a hypothetical individual who stutters (during the baseline condition), indicating that imagining stuttering and actually witnessing it can elicit similar affective responses. This notion can be accounted for by mirror neuron activation and their relation to ‘theory of mind’, whereby these neurons allow those who observe actions (i.e. stuttering) to place themselves in the ‘mind of others’ performing those actions. Those who have cared for autistic or severely aphasic patients for relatively short periods of time have the ability to interpret correctly small amounts of ‘garbled’ information with respect to their daily living needs. Most likely, all humans are biologically predisposed to play the roles of both sender and receiver in the communicative process and can therefore identify a breakdown in either role. Most people have probably witnessed stuttering at some point in their lives and can empathize with the unpleasantness surrounding it. Hence, asking participants to hypothetically consider stuttering was seemingly sufficient to elicit negative responses that were similar to those recorded when actually witnessing it.

Findings also showed no difference in mean ratings between first versus second stuttered and first versus second fluent speech samples on all 11 rating scales, suggesting that participants’ affective responses to the second stuttered or fluent speech sample did not change after viewing the first stuttered or fluent speech sample for a period of 30s. These findings indicate a strong impact of stuttering on a listener’s emotional responses. Further, on the dimensions of happy–unhappy, pleasant–unpleasant, avoiding–approaching, embarrassing–soothing, and annoyed–joyful, a significant difference was observed between the baseline affective responses to hypothetical individual who stutter (before viewing experimental stimuli) and combined affective ratings after viewing the first and second stuttered speech samples. The difference in a listener’s self-ratings on the above rating scales may be attributed to the task itself. Simply put, during the baseline, participants rated their responses to a general hypothetical person who stutters as compared to the specific stimuli (i.e. speech samples of severe stuttering) during the experimental task. On the other hand, the data showed that participants were calm, relaxed, and joyful while observing the fluent speech samples presented in this study compared to observing fluent speakers in general. Although there may be individual differences among participants’ affective responses to stuttered and fluent speech samples, the present findings are indicative of severe stuttered speech eliciting a set of negative emotional responses in listeners.

However, further investigation is warranted to understand the specific effects of degree of severity of stuttering, gender, and visual and audio only presentation of stimuli on the affective and physiological responses and emotions of listener’s per se. In other words, specific effects to the auditory or visual behaviours of stuttering on a listener’s emotional and physiological responses are not discernible from the results of the present study. We suspect that both primary (i.e. repetitions and prolongations) and secondary (i.e. aberrant facial manifestations) behaviours of stuttering are compensatory in nature to release the neural block and will have a powerful impact on the listener’s physiological and emotional state. However, further research is needed to empirically evaluate this notion and understand the
listener’s psychophysiological responses to a person who stutters with and without any aberrant facial and non-auditory behaviours.

The results strongly indicate that normally fluent individuals report negative affective responses when presented with stuttered speech as compared to fluent speech. These findings could be useful in therapeutic counselling as the listener’s affective and visceral responses to stuttered speech can be explained to people who stutter. Clinicians should be acutely aware of this reality and encouraged to share this information with their patients. Being cognisant of the affective responses that stuttering can elicit in others may potentially help those who stutter develop appropriate coping strategies. These findings provide empirical evidence that contradicts the beliefs of some clinicians who counsel stuttering clients (in order to improve their morale) that stuttering is one minor problem and persons who stutter are not perceived to be different from persons who do not stutter (Kalinowski et al. 1996). In contrast, we as clinicians and caregivers must learn to recognize the affective responses and help educate those who stutter as to their involuntary nature. Some of these autonomic responses are reflexive and primitive in nature. Because there is little hope of changing the typical autonomic responses to stuttering, we advise those who stutter to join support groups such as British Stammering Association, National Stuttering Association, Stuttering Foundation of America so that they can share their similar emotional experiences with others who stutter. These findings may also explain the genesis of general public attitudes and perceptions about people who stutter. We suggest that as listeners’ begin to feel the negative emotions associated with the breakdown in communication and aberrant speech behaviours, their emotional state is sensed and transferred via mirror system to the person who stutters. That is, as listeners report being emotionally aroused, unhappy, nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassed, and annoyed while in the presence of stuttering, they could be imparting some or all of these feelings upon the person who is stuttering.

In summary, these findings suggest that viewing aberrant speech and facial movements associated with severe stuttering elicits affective arousal and the feelings of being unhappy, nervous, uncomfortable, sad, tense, unpleasant, avoiding, embarrassing, and annoying in listeners compared to observing of normal fluent speech. This is the first study to report such changes in affective responses to stuttered speech using the SAM rating scale (Bradley and Lang 1994). In agreement with the self-ratings, listeners’ showed increased skin conductance and reduced heart rate while viewing the stuttered speech relative to the fluent speech. These findings may help clinicians become acutely aware of listener’s physiological responses and affective attitudes to stuttered speech. They may also help patients develop coping strategies to deal with these responses and attitudes. Further, the present findings help provide a plausible explanation for the general public perceptions about people who stutter.

References


Responses of fluent listeners while watching the speech of adults who stutter


### Appendix A

The following questionnaire was used to evaluate a participant’s emotional responses to stuttered and fluent speech samples:

*Please rate your feelings to the stimulus on the following scales.*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Rating (1-9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nervous</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>2. Comfortable</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>3. Funny</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>4. Tense</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>5. Pleasant</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>6. Avoiding</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>7. Emotional</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>8. Embarrassing</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>9. Annoyed</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
</tbody>
</table>

Responses of fluent listeners while watching the speech of adults who stutter