THE INTERNATIONAL JOURNAL OF HUMANITIES & SOCIAL STUDIES

Investigating Stress-Related Speech Features Effect of Perceived Stress, Mindfulness and Personality on Stress-Related Speech Features

Apurvi Khaitan
Ex-Student, Department of Psychology, University College London, England

Abstract:
What is the role of perceived stress, mindfulness and personality in predicting stress through acoustic features? This study aims to explore these three psychological dimensions by asking 86 participants to record their speech over 10 days using their smartphones. Based on their relation with stress, the study hypothesizes that perceived stress and mindfulness would predict stress-related speech features whilst personality dimensions would explain the variance in features. Linear mixed effect found both perceived stress and mindfulness to predict increase in fundamental frequency, while certain personality dimensions found increase in F0 and Shimmer. The findings suggest the possibility of using acoustic features as a measure of stress.

1. Introduction

Stress is a psycho-physiological response to tackle the changes and challenges in our lives (Yao, Jitsuhiro, Miyajima, Kitaoka& Takeda, 2015). As it carries a wide range of health-related diseases, it has received a great attention in well-being research. Hence, in order to help individuals, manage this problem, it is essential to optimize technologies that record stress and further our insights on factors that influence it.

Research has documented a variety of ways to measure stress, this current study aims to investigate psychological stress using a speech-based technique and collect the acoustic samples using smartphones. Factors that have found to influence stress such as mindfulness and personality traits are also explored. Given the extent to which one's stress levels is dependent on one's perception and appraisal; self-reported stress is also examined. The review begins by providing a general understanding of stress and its measures. It is followed by discussing stress assessed through speech. Consequently, the relation of mindfulness and personality factors with stress are explained and lastly, the theatrical rationale of the current study is presented.

1.1. Understanding Stress

The term 'stress' can be interpreted in several ways: a) In our day-to-day conversations, stress occurs when we are unable to perform tasks in the time available to us;b) engineers/physicists describe stress as a force that produces a change in shape or volume c) in the psycho-physiological sense, stress is viewed as an aspect of the environment that causes mental, emotional and physical strain within individuals (Collins English Dictionary, 2010). In absence of a single definition of stress applicable to all disciplines, we focus on the psycho-physiological aspect of stress.

The *physiological response* to stress is perceived as the body's attempt to regain homeostatic balance following an internal or external threat (Kirchhubel, Howard &Stedmon, 2011). A stressful reaction causes the activation of the hypothalamic-pituitary-adrenal axis (HPAA) that leads to the release of the two main stress hormones; glucocorticoids (cortisol) and catecholamines (norepinephrine) (Cohen, Kessler & Gordon, 1995). The secretion of these give rise to the 'fight-or-flight syndrome' that alerts the body to tackle an emergency situation (Canon, 1915). As a result, the body experiences an increase in heart rate, respiration rate and muscle tension, alongside pupil dilation and reduced salivation (Jessen, 2006). Alternatively, according to Mason (1968), there are three *psychological determinants* that induce stress within individuals. It requires the individual to interpret the situation as novel, unpredictable and feeling a loss of control over the situation. Psychological stress manifests as negative affective states making the individual feel discomfort, anxiety and pressure. Now that we are aware of both the physiological and psychological variables that indicate stress within individuals, we go on to examine the measurement of stress.

1.2. Measuring Stress

Both cortisol and catecholamines are easily measurable in blood and urine, however over the past 20 years' salivary cortisol became the most popular biomarker in stress studies (Kirschbaum & Hellhammer, 1989).This became the preferred technique as it did not require skilled personnel to install a catheter and extract the hormone; making it more convenient. Another method that gained popularity was the measurement of cortisol in hair: it provided the accumulation of cortisol over time (1-2 months), unlike blood or saliva that only captured daily fluctuations in cortisol levels (Russell, Koren, Rieder, & Van Uum, 2012).

However, due to their impracticalities, these methods may not be viable. First, monitoring cortisol through blood, urine and saliva requires multiple samples over a prolonged period of time to estimate the cortisol levels (Kudielka & Kirschbaum, 2005). For example, the laborious task of collecting cortisol in urine requires 24 hours and only reflects the previous day's cortisol level (SauveWalsh, Tokmakejian & Uum, 2007). Secondly, due to circadian effects, the samples for blood and saliva need to be taken at specific times of the day (Yang et al., 2001). Lastly, assessing cortisol in hair is restricted to individuals with sufficient hair at the posterior vertex; Moreover, it is a reliable measure for chronic stress though, less so for acute stress (Sauve et al., 2007). In general, quantifying stress through its biomarkers appears to be time-consuming and an invasive procedure, if used outside of medical contexts. On the other hand, self-report measures are increasingly used to measure stress in its psychological form. An example of this is the Perceived Stress Scale (PSS; Cohen, Kamarck & Mermelstein, 1983) that has been well validated and effective in mindfulness-based interventions to combat stress (Carmody& Baer, 2007).

1.3. Speech under Stress

Another interesting way of measuring stress comes from investigating its effect on non-linguistic features of speech. Research has shown that apart from the linguistic content, a wealth of information resides in the acoustic features of the speaker from which the listener could derive his emotional state and intentions (Protopapas& Lieberman, 1997). It is believed that the physiological alterations that accompany a stress-reaction cause changes in the acoustic speech signal. For example, when stressed, the body adapts to the condition through the tensing of muscles in the vocal folds or changing of the shape of the tract (Yao et al., 2015); there is also a rise in respiration leading to increased airflow(Kurniawan, Maslov, & Pechenizkiy, 2013). Thus, the variations in the physiology of the vocal system and in aerodynamics produces stressed speech.

Scherer (1986) proposed a "sequence theory of emotional differentiation" to explain the change in the physiology of speech production system in relation to different emotional contexts. According to the theory, stimuli are evaluated based on certain functional parameters such as "need/goal", "urgency" and "coping potential". The outcomes of the evaluation affect the nervous system which in turn defines the voice characteristics that carry the emotional information. For example, unpleasant stimuli cause "faucal and pharyngeal constriction and tensing along with shortening of the vocal tract"; leading to stronger high-frequency resonance.

The application of voice-based stress detection spans across multiple fields: a forensic psychologist could assess the mental state and amount of stress experienced by hostages/criminals by simply listening to their voice; a psychiatrist could determine the risk of suicide of his/her patients; ground personnel and air controllers could evaluate the extent of danger in flight based on the pilot's speech (Kirchhubel, Howard & Stedmon, 2011). As a result, measuring stress through speech is an ideal well-established method. One of its key advantages is that it is non-invasive as it does not require medical personnel or the presence of traditional monitoring devices which is inherently more stressful and prevents objective examination. Secondly, due to technological advancements, smartphone applications are able to collect speech input allowing the patient to be in their natural environment. This makes data collection easy and accessible. Furthermore, stress levels can be automatically identified by the user's smartphone, enabling the individual to recognize and participate in stress-reduction activities.

Researchers have manipulated stress in various contexts such as through physical exercise, cognitive workload, stress inducing situations and laboratory conditions to gain insights on its relationship with speech. For example, physical task stress through strenuous activities result in changes in the speech production systemand in turn, speaking behaviour. The underlying reason is that both speaking and exercise compete for the same resource, i.e. ventilation; which is significantly altered during and between utterances whilst exercising. As a result, there is a corresponding impact on muscle control and airflow within the glottal excitation structure and vocal articulatory tract (Godin & Hansen, 2015). In general, speech during exercise causes reduction in oxygen intake, an increase in blood lactic acid, a decrease in physical performance and hastens fatigue.

Similarly, researchers investigated the effects of cognitive workload on speech production. Lively, Pisoni, Summers and Bernacki (1993) conducted a study in which workload was induced by asking participants to perform a compensatory visual task whilst speaking aloud test sentences. Changes in acoustics were measured by comparing utterances under workload and control conditions. The findings indicated laryngeal and sub laryngeal adjustments as well as modifications in the timings of articulatory gestures. Again, performance suffers because individuals have limited capacity of attentional resources and both speaking and cognitive tasks require attention. Thus, stress caused by cognitive workload and physical strain alter speech and its effects on acoustic features is later explained.

It can be understood that stress impacts speech in various ways, as addressed at the ESCO-NATO workshop on 'Speech under Stress", in an effort to consolidate the concept of stress in speech contexts (Trancoso& Moore, 1995). The biggest challenge faced by them was in conceptualizing stress, since they realized that stress manifests in various dimensions. As a result, they drew up a taxonomy of stressors with four categories and corresponding stressors; Physical (physical strain, illumination), Chemical (alcohol, chemical drugs), Physiological (injuries, disease) and Psychological (emotion, workload). While there are various ways in which stress can be induced in speech, this study focuses on the relation between *psychological stress* and speech features.

1.4. Acoustic features

Several acoustic features of stress (F0, Intensity, Jitter and Shimmer) have been identified as robust correlates of psychological stress; all of which are explained in further detail.

- Fundamental frequency (F_0). $F_0$is defined as the rate of vocal fold vibration and is the most investigated parameter when exploring stress in speech. It is the parameter that carries both semantic and the emotional state of the utterance, and contributes to the perceived pitch of the sound produced. The physiological change that causes higher fundamental frequency is attributed to an increase in muscle tension resulting in the tensing of vocal folds that occurs during stress (Kirchhubel, Howard & Stedmon, 2011). It has been found that age (Harrington, Palethorpe& Watson, 2007) and gender (Baken & Orlikoff, 2000) affect the mean F0. Previous studies have induced stress in various ways such as cognitive overload (Lively et al., 1993), situational stress (Hicks, 1979), real life stress (e.g. of pilots) (Huttunen, Keranen, Vayrynen, Paakkonen & Leino, 2011) and found an increase in F0 as stress increases. It has also been demonstrated that F0 increases in real-life stress (e.g., public speaking condition) as opposed to laboratory conditions where a non-significant increase was found (e.g. electric shocks) (Hicks, 1979).

- Intensity. Intensity refers to the loudness of speech, an increase of which has been found to predict greater stress. During a stress reaction, the rise in respiration rate causes an increase in the sub-glottal pressure and airflow in the vocal tract that leads to a higher intensity (Kirchhubel, Howard & Stedmon, 2011). Tiffin and Steer (1937) found an average intensity increase of 9.3 dB between stressed and unstressed words. Moreover, on a syllabic level, Ortleb (1937) found that emphasized syllables in cases of emotional and unemotional materials were more intense than unstressed syllables, with more pronounced intensity for emotional utterances. The relationship between stress and intensity has not been extremely evident with some showing an increase in intensity (Brenner & Shipp, 1988) whilst others were inconclusive (Hecker et al, 1968).

Initially, the overall amplitude was measured to gage its link with stress. In general, most studies have found increase in intensity when stress was induced through cognitive overload (Mendoza & Carballo, 1998, Lively et al., 1993). However, recent studies have started to investigate the amplitude of different frequency bands as it is believed that intensity measurements are greater at higher frequency bands, i.e., above 1000 Hz in stressed conditions verses control conditions (Kirchhubel, Howard &Stedmon, 2011). As a result, root mean square is considered to be a more accurate measure of intensity as compared to the peak amplitude of signal energy.

- Jitter and Shimmer. Jitter refers to cycle-to-cycle variations in vocal fold vibration. Scherer (1979) asserts that jitter could be caused by lack of coordination between the laryngeal muscles that contribute to phonation and disrupted respiration rate that occur during stress. The findings with regard to jitter are inconsistent with a few reporting a decrease (Lively et al., 1993). Conversely, shimmer known as the cycle-to-cycle variability in amplitude measurements, has been given less importance. Reduced shimmer was observed by Mendoza and Carballo (1998) whilst a significant increase was found by Lively et al. (1993) in relation to stress, hence the findings have been inconclusive.

1.5. Mindfulness and Stress

During their lifetime, individuals face various kinds of demands and challenges that could have an adverse impact either physically or psychologically. Yet there are notable differences in the way people react to these difficult life-events which in turn affect their well-being (Larsen, 2000). In the recent years, there has been considerable interest in mindfulness and its programs because they are believed to be a protective factor when dealing with these difficult situations. Thus, it is crucial to understand mindfulness and how it aids well-being.

Mindfulness is defined as the state of attention and awareness that allows individuals to be conscious of occurrences at present (Brown & Ryan, 2003). The key characteristics of mindfulness are the monitoring and acceptance of those experiences at present (Quaglia, Brown, Lindsay, Creswell & Goodman, 2014). According to Ryan (2003), since individuals possess the conscious ability to attend and be aware of their cognition, individual differences exist in the capacities of mindfulness; this is relevant to stress as varied propensities of it can modulate stress responses.

The Mindfulness Stress Buffering Account proposed by Creswell and Lindsay (2014) elucidates how mindfulness reduces stress appraisal and minimizes stress-reactivity responses. Firstly, the theory posits that mindfulness increases the *recruitment of stress-regulatory regions of prefrontal cortex* that inhibits the activity in stress processing regions. Secondly, mindfulness also decreases the reactivity of the central stress processing regions (e.g., amygdala, anterior cingulate cortex, hypothalamus) that are responsible for signaling the HPA axis. Moreover, the Buffering account also predicts that mindfulness interventions are more pronounced for those who have high stress burdens (e.g., unemployed adults) as opposed to those with low-stress. This hypothesis was supported by Brown et al., (2012) who found higher levels of mindfulness and lower cortisol reactivity among participants in high stress conditions (perform math and speech tasks in front of evaluators), whereas it found no association between the two in the low stress conditions (perform the same task alone). Thus, the account provides a comprehensive link between stress and mindfulness.

Mindfulness-based stress reduction (MBSR, Kabat-Zinn, 1982) is the most widely used mindfulness intervention that has been increasingly recognized as a solution to a variety problems and disorders. MBSR sessions involve three primary meditation exercises namely body scan, sitting meditation and gentle yoga; all of which encourage a non-judgemental stance of one's bodily sensations and emotional states. The intervention has shown promising results with improvements in HPAA responses and decrease in salivary cortisol levels (Marcus et al., 2003). The role of mindfulness with perceived stress reduction has also been documented across a large number of studies and across different populations (Carmody& Baer, 2007). These findings further establish the association between stress and mindfulness.

In the attempt to identify key facets of mindfulness, Baer, Smith, Hopkins, Krietemeyer and Toney (2006) conducted an exploratory factor analysis using five independently developed

Facet	Description
Observing	Notice or attend to internal and external experiences.
Describing	Use words to explain internal feelings.
Acting with awareness	Attend to events of the present.
Non-judging of inner experience	Take a non-evaluative stance towards thoughts and emotions.
Non-reactivity of inner experience	Allow thoughts and feelings to come and go without getting caught up

Table 1: Description of the FFMQ Facets

developed scales. They concluded a five-factor structure showcasing various dimensions of mindfulness, referred to as the Five Facet Mindfulness Questionnaire (Carmody& Baer, 2008). The facets are explained in Table 1

As mindfulness reduces an individual's reactiveness towards stress, it can be argued that these facets too may have individual correlations with stress. For example, *Non-reactivity of inner experience* subscale instills a nonchalant attitude within an individual as compared to *Observing* subscale which makes an individual more aware of their responsibilities. Hence, it is plausible that people might be relatively more stressed in the latter subscale due to their varied thinking. However, the relation between the dimensions of mindfulness and perceived stress has not received a great deal of attention; presenting an opportunity to study this relation.Branstrom, Duncan and Moskowitz (2011) found medium negative correlations (-.29 to -.58) between the last four FFMQ subscales and the PSS though, mindfulness facets have not yet been fully evaluated with reference to perceived stress.We know that differing levels of mindfulness across individuals may have a considerable impact the level of stress experienced. Therefore, it is of great interest to explore the extent to which mindfulness impacts one's perception of stress.

1.6. Personality and Stress

Personality impacts a person's thinking, behaviour and feelings. (George, 1992). Research has shown that personality and stress are closely related as personality is considered to be crucial for the selection and shaping of stressful events as well as coping with stress. Hence, it is important to explore the different studies and theories that show the effect of personality on stress levels.

The Five Factor Model (FFM) measures independent personality factors such as Neuroticism, Extroversion, Agreeableness, Openness and Conscientiousness (Costa & McCrae, 1992). The Neuroticism subscale that is comprised of anxiety, depression and vulnerability is found to significantly predict anxiety while adversely affecting stress. Individuals that are highly neurotic consider banal everyday situations to be threatening. Thus, studies reflect a high degree of neuroticism to be correlated with occurrences of negative events later in life (Vollrath, 2001). Investigating the associations between the FFM and the PSS has shown that higher conscientiousness was related to lower PSS (Besser & Shackelford, 2007) while no relation existed between PSS and Extroversion or Agreeableness (Grant &Langan-Fox, 2007). This demonstrates that personality effects the appraisal of stress and that the primary driver of PSS is neuroticism (Conard& Matthews, 2008).

Another area of research considers the physiological reactions to stress that differs across personality dimensions; concentrating on changes in heart rate, blood pressure and the excretion of stress hormones. Researchers came up with two types of personalities A and B (Spector & O Connell, 1994). Urgency, competitiveness, hostility and dispositional anger are characteristics of Type A personalities, and were found to deal with stress aggressively. Findings showed that hostility and irritability are more related to stress than other traits (Matthews, 1988). Moreover, Type A personalities were more likely to react with increased cortisol and catecholamine, indicating stress and increased risk of coronary heart diseases. Conversely, Type B personalities are easy going and more accepting of emotions. They are supportive of others, express positive feelings and are more satisfied with their jobs (Kirkcaldy et al., 2002. Hence, they are far less prone to be stressed than Type A individuals.

Recently, type D, a new concept indicating 'distressed' personality, has been introduced (Denollet, 2000). These individuals obtain high scores on negative affectivity and social inhibition personality dimensions. A plethora of studies have attended to the relation between negative affectivity and cortisol activity through structured laboratory stressors; public speaking and mental arithmetic and aversive stimulation, and have found a significant positive correlation between the two (e.g., Holsboer, 2000; Sher, 2004).

Furthermore, the examination of stress-producing mechanisms, leads to the understanding of the effects of certain personality traits on the origins of an individual's stress. (Bermudez,1999). These mechanisms affect one's personal goals, needs and strivings; thus influencing the level of stress. An example of such a mechanism is the case of illusionary goals; individuals who chase after unrealistic goals, look for situations that require more than their abilities that leads to increased probability of stress (Semmer, 1996). This particular mechanism presents a degree of danger for individuals with high levels of Self-efficacy; who believe they are masters at problem solving (Bandura, 1992).

The risk behaviours mechanism (Bermudez, 1999) focuses on individuals with a high Extraversion and Sensation-seeking (e.g. reckless driving and the consumption of drugs) who are more likely to partake in high-risk activities (Headey & Wearing, 1989). Studies have shown that these personality traits eventually result in high stress. Moreover, high Psychoticism and low Conscientiousness are related to health risk behaviours (e.g. smoking and spontaneous sexual activity) which eventually lead to stress (Booth-Kewley& Vickers, 1994) Therefore, it is evident that different personalities interact with stress in various ways.

Lastly, the association between personality and stress can also be observed in daily activities. Ge et al., (2014) measured four driving behaviours; negative and cognitive/emotional driving, aggressive driving, risky driving, and drunk driving. Anger was positively correlated with the former three driving behaviours whilst altruism showed a negative correlation with aggressive and drunk driving. Overall, the results depicted stress as an important factor influencing driving behavior, mediated by the effect of personality traits. Dumitru & Cozman, (2010) concluded that nurses with high stress levels had lower empathy, independence and tolerance. These results echoe the influence of personality in determining stress, and the importance of investigating this relationship. The Personality Inventory of DSM-5 (PID-5) is a reliable self-report measure that is used to assess personality traits and psychopathy (Krueger, Derringer, Markon, Watson, &Skodol, 2012).

1.7. Current Study and Hypotheses

Although measuring stress through speech is a widely well-established technique, there is a surprising lack of research on the use of self-report stress questionnaires to examine relationship. Unlike objective measures, self-report measures account for the inter-personal differences that occur when interacting and appraising stressful situations which in turn lead to differing levels of perceived stress. Factors such as one's daily events, habituation and coping styles influence perceived stress which self-reports are able to account for (Cohen et al., 1983; Kirschbaum et al., 1998). Moreover, since salivary cortisol is highly correlated with PSS, it would be interesting to investigate the relation between PSS and speech features; as an alternate objective measure of stress.

A number of studies have considered the analysis of speech under both simulated and actual stress conditions. However, the ecological validity of previous stress experiments has been questioned. It has been found that simulated stress (e.g. via electric shocks) (Hicks, 1979) produces highly exaggerated misrepresentations of emotions in speech. In addition, previous studies that measured stress in natural settings (e.g. of pilots and criminals) have reduced scope for generalization, as the stress levels observed in those environments cannot not mirrored in day-to-day activities. Contrastingly, this study examined the acoustic correlates within individuals, through multiple speech recordings, over a period of 10 days. This methodology has three main benefits: a) it adds ecological credibility to this field that mostly monitored induced stress b) the average of the recordings will produce a more reliable estimate of the acoustic features of stress c) by observing natural stress, it will contest the criticism that these acoustic correlates are artefacts of experimental designs. Moreover, collecting data via smartphones will enable a greater sample size; which will help adhere to previous findings with higher validity and improved statistical power.

Several methodological limitations in the study of speech under stress were also addressed. Previous studies have computed peak amplitudes as a measure of speech intensity, however, the mean root mean square (rms) energy may provide a better estimate of the average intensity across the entire speech sample, without unrelated disturbances such as throat clearing or coughs having a significant influence. Furthermore, this study takes into consideration high-level functional of speech features (e.g. maximum); giving a greater insight into how speech is modified by the perception of stress.

Furthering, MBSR interventions have shown that greater mindfulness leads to a reduction in reactivity to stress (measured through salivary cortisol). However, no previous studies have examined the facets of mindfulness in relation to perceived stress. It is known that Describe and Observe subscales are associated with the perception of stress itself, hence it was hypothesized that scores on the Describe and Observe subscales would predict an increase in stressful speech features. On the other hand, acting with Awareness, Non-judging and Non-reacting subscales are associated with internalizing and reacting to stress and is hypothesized to predict a decrease in stress linked acoustic features.

Similarly, based on the personality dimensions elucidated by FFM and Tupe A/B model, findings have shown that stress is vastly mediated by these varied traits. This is because these traits affect appraisal and coping mechanisms which play a crucial role in levels of stress. As a result, exploring these personality traits in relation to stress-related speech features will further our insights on the relationship.

The study proposed a mixed model design that would examine acoustic features across multiple recordings with the three sale-reports. This study aims to investigate stress related speech features in relation to perceived stress; mindfulness and personality traits. Hence, it proposes to corroborate previous findings that show increase F0 and intensity, decreased jitter, during stress conditions.

The following are the hypotheses:

- H1.Higher perceived stress will predict greater stress-related speech features (increased F0 and intensity; decreased jitter).
- H2.Individual dimensions of mindfulness will show different associations with perceived stress and its speech features. Specifically, the Describe and Observe facets will positively correlate whilst the Act with Awareness, Non-judging and Non-reacting facets will negatively correlate with perceived stress and its speech features.
- H3.Shimmer will not be predicted by or significantly correlated with PSS or FFMQ scores.
- H4.In an exploratory way, individual differences in personality will explain the variance in stress-related speech features.

2. Method

2.1. Participants

This study recruited 210 participants and utilized a number of platforms such as social media, University College London Psychology subject pool and word of mouth. The final sample comprised of 86 participants (29 males and 57 females) with ages ranging from 18 to 58 (M = 25.55 and SD = 10.14). Thirty-eight of 86 (44%) participants were native English speakers. The remaining 124 participants were excluded as they either did not complete the self-report questionnaires or submitted less than 10 recordings. Those who had successfully completed the study, were entered into a prize draw to win one of the 5 £50 Amazon vouchers. The study was approved by the UCL Graduate School Research Ethics Committee (ID: 6129/001).

2.2. Apparatus and Stimuli

The information sheet, instructions to the study, consent form and the self-report questionnaires were administered via an online survey platform (http://speech.soma-analytics.com/start/) that could be accessed by participants using their preferred web browsers. The participant had to install a mobile application called “MoMo by SONA” onto their smartphones across iOS and Android platforms. This application presented the users with the stimuli set of 20 validated neutral English sentences and recorded their speech using the smartphone’s inbuilt microphone. Our collaboration with SOMA Analytics (http://www.soma-analytics.de) helped us conduct data collection, application design and self-report surveys.

2.3. Measures

- Perceived Stress Scale (PSS). The PSS-10 (Cohen, Kamarck & Mermelstein, 1983) is the most widely used psychological instrument to measure the extent to which situations in one’s life are appraised as stressful over the past month. It is a 10-item questionnaire and each item is rated on a 5-point scale starting with (0) never; (1) almost never; (2) sometimes; (3) fairly often; (4) very often. The total score is obtained by reversing the scores on the four positively stated items (4, 5, 7 and 8) and then summing across all 10 items. The scale has a potential score of 0-40 with a score of 13 considered as average, and scores exceeding 20, indicative of high stress. The PSS-10 has shown good internal consistency (α=.82) with high test-retest reliability and their use in health care practice is warranted (Andreou et al., 2011). PSS scores are positively correlated with high blood pressure and cortisol levels, failure to quit smoking, depression and social anxiety (Cohen & Williamson, 1988).
- Five Facet Mindfulness Questionnaire (FFMQ). The FFMQ (Baer, Smith, Hopkins, Krietemeyer& Toney, 2006) is a measure that aims to combine five independently developed mindfulness questionnaires in an integrative manner. It consists of 39-items that represent the five elements of mindfulness such as Describe, Observe, acting with awareness, Non-judging of inner experience and Non-reacting to inner experience. Items are scored on a 5-point Likert-type scale starting with (1) never or very rarely true; (2) rarely true; (3) sometimes true; (4) often true; (5) very often or always true. Facet scores are computed by summing the scores on the individual items with scores ranging from 8 to 40 (except for the non-reactivity facet, which ranges from 7 to 35). FFMQ has a score range of 39-195 with higher scores indicating increased levels of mindfulness. Good internal consistencies for all five subscales have been reported with Cronbach’s alpha ranging from .72 to .92 (Baer et al., 2008).
- The Personality Inventory for DSM-5 (PID). The PID-5 (Krueger et al., 2012) is composed of 220-items. However, this experiment adopted the reduced 100-items version to save the time of participant’s in answering the questions. The self-report aims to measure a variety of personality dimensions and the items are scored on a 4-point Likert scale ranging from 0=Very False or Often False to 3= Very True or Often True. The questionnaire has five broad dimensions and each of these have subscales within them: *Negative affect* (Anxiousness, Emotional lability, Hostility, Preservation, Restricted affectivity, Separation insecurity and Submissiveness)*, Detachment* (Anhedonia, Depressivity, Intimacy avoidance, Suspiciousness and Withdrawal), *Psychoticism* (Eccentricity, Perceptual dysregulation and Unusual beliefs and experiences*), Antagonism* (Attention seeking, Callousness, Deceitfulness, Grandiosity and Manipulativeness) *and Disinhibition* (Distractibility, Impulsivity, Irresponsibility, Rigid perfectionism and Risk taking).Scores for the dimensions were computed by calculating the total from each of their subscales and then averaging (scores of Restricted affectivity and Rigid perfectionism were to be reversed). The Cronbach’s alpha for the reduced version range from .89 to .91 with facet scores ranging from .74 to .88, thus both showcasing excellent internal consistency.

2.4. Design and Procedure

A mixed design was adopted using between-subjects variables as perceived stress, mindfulness, personality and acoustic features, and within-subjects as the multiple audio recordings. The independent variables consisted of the questions asked in all three self-reports as well as the stimuli of English sentences. The dependent variables were the scores obtained on each of the questionnaires as well as the non-linguistic speech features that were extracted of each participant.

The study began with participants receiving an email with a link to the study’s sign up webpage. Participants were first presented with instructions and information about the study and were asked for their consent to take part in the study. Following which, the participants were required to enter their demographic details and complete PSS, FFMQ and PID questionnaires. After answering the questionnaires, the participants were instructed to download the study application called “MoMo by SONA” through their smartphone’s app store. Each user was given a unique token that was to be entered at the start of the application; allowing them to access the study. The token also served as the participant’s identifier for data analysis. The recording task presented an English sentence on the screen to be read aloud through the smartphone’s microphone. The acoustic features of the recordings were analyzed using an algorithm upon the submission of audio files. Participants were encouraged to speak loudly and clearly and to make their recordings in a relatively quiet environment. A success message would appear after each recording indicating that it was correctly submitted. The participants were required to complete 3 recordings per day over 10 consecutive days. They were sent push notifications at the hours of 08:00, 14:00 and 20:00 as reminders to complete the voice recording task.

2.5. Data Reduction

The openSMILE toolkit (v2.1.0, http://www.audeering.com/research/opensmile; for details, see Eyben, Weninger, Gross & Schuller, 2013) was used to extract the acoustic features from the audio files. The algorithim decomposed each audio file into 25-32ms frames with four low-level features being extracted from each frame: i) root mean square (rms) energy; ii) F_0; iii) jitter, and iv) shimmer. Six high-level functionals including maximum, mean, minimum, standard deviation, skewness, and kurtosis were calculated for each low-level feature; of which the first two were used in the statistical analysis. Other details on the formulae and extraction of these features can be found in Appendix B.

2.6. Statistical Analysis

We first explored the data by finding the descriptives of the self-report measures and speech features. Next, we assessed the internal consistencies within each of the subscales of PSS-10, FFMQ and PID-5. Bivariate correlations across all speech samples were conducted to infer the associations between speech features with respect to perceived stress (hypothesis 1), mindfulness (hypothesis 2) and personality traits (hypothesis 3). Based on the significant relationships that were identified, analysis employing the linear mixed effects model were performed with the PSS, FFMQ and PID as fixed predictors of speech features, with scores nested within participants' multiple recordings. As there is no standard measure of the effect of size in the linear mixed effects models, an estimate of Cohen's *d* was used. Bivariate correlations of PSS with respect to FFMQ and PID were also studied to gage their relation with self-reported stress as opposed to the objective measurement via speech. Lastly, age and gender effects on the acoustic features were investigated to explain variance in observed stress.

3. Results

3.1. Descriptive Statistics

The final analysis was conducted from the data of 86 participants as they completed the study in its entirety by answering all questionnaires and had more than 10 recordings. They completed between 10 and 39 recordings (*M*= 22.91, *SD*= 7.46), providing a total of 1970 recordings. As mentioned, the data analysis focuses on the maxima and means of each of the four low-level features. Descriptive statistics for the psychological questionnaires as well as the acoustic features are shown in Table 2 and Table 3 respectively.

Acoustic Feature	*N*	*Mean*	*SD*
Intensity (max)	1970	.044	.044
Intensity (mean)	1970	.005	.005
F0 (max)	1970	332.597	93.683
F0 (mean)	1970	43.532	41.152
Jitter (max)	1970	.699	.186
Jitter (mean)	1970	.051	.022
Shimmer (max)	1970	.890	.139
Shimmer (mean)	1970	.129	.036

Table 2: Means and Standard Deviations (SD) for each of the extracted speech features
Note: N= number of recordings; F0= Fundamental frequency

3.2. Internal Consistencies and Inter-correlations

Cronbach's alpha statistics were computed for the PSS; illustrating excellent internal consistency, and FFMQ and PID; showcasing good internal consistencies for each of the subscales. The Disinhibition subscale of PID, showed moderate internal reliability (α=.52). These are shown in Table 3

Scores on the PSS and FFMQ were significantly correlated with the speech features which can be found in Table 4. PSS scores significantly positively correlated with F0 mean and negatively with jitter and shimmer. A significant negative association was found between Observe subscale and intensity mean and F0 mean. Contrastingly, describe subscale did not have significant correlations with any of the speech features. Acting with awareness and Non-judging had significant negative correlations with F0 mean and intensity (max) respectively. Non-judging was found to have positively significant correlations with jitter and shimmer and lastly, Overall scores had significant negative correlations with F0 and positive with jitter.

Scale	Cronbach's *α*	*N*	*Mean*	*SD*
PSS	.88	86	18.73	6.57
FFMQ				
Observing	.62	86	24.40	5.21
Describing	.64	86	27.70	5.90
Acting with awareness	.66	86	25.69	5.77
Non-judging	.70	86	24.44	6.37
Non-reacting	.74	86	21.22	4.28
Overall	.86	86	123.45	15.82
PID				
Negative affectivity	.69	86	4.20	2.50
Detachment	.72	86	2.93	1.90
Psychoticism	.71	86	2.72	1.99
Antagonism	.78	86	2.83	1.94
Disinhibition	.53	86	3.60	2.13

Table 3: Cronbach's alpha, mean scores and Standard Deviations (SD) for the psychological questionnaires
Note: PSS=Perceived Stress Scale; FFMQ=Five Facet Mindfulness Questionnaire; PID=The Personality Inventory of DSM-5

Moreover, scores on the PSS were all negatively significantly correlated with FFMQ scores, with the Observe scale at *r* (84) = -.110, *p*<.001; with Describe scale at *r* (84) = -.215, *p*<.001; with Acting with awareness scale at *r*(84)= -.496, *p*<.001; with Non-judging scale at *r*(84)= -.423, *p*<.001; with Non-reacting scale at *r*(84)= -.535, *p*<.001 and FFMQ Overall at *r*(84)= -.611, *p*<.001.
Scores on the PID were also significantly correlated with the acoustic features found in Table 5. Negative affectivity was found to significantly positively correlate with intensity mean and F0 mean. Detachment significantly correlated with all four speech features, positively with F0 mean and negatively with intensity mean, jitter mean and max and shimmer mean. Psychoticism showed positive significant correlations with F0 mean and max as well as shimmer mean. Antagonism was found to significantly correlate with intensity mean and shimmer mean in a positive manner. Finally, Disinhibition had positive significant correlations with shimmer mean.

Speech Feature	**PSS**	**FFMQ**					
		O	D	AA	NJ	NR	Overall
Intensity (max)	.06**	-0.04	0.03	-0.02	-.07***	-0.03	-.05*
Intensity (mean)	0.02	-.08***	-0.007	0.01	-0.03	-0.006	-0.04
F0 (max)	-0.002	-0.03	-0.03	-.07**	0.04	0.002	0.02
F0 (mean)	.18***	-.08***	0.02	-.15***	0.002	-0.03	-.08***
Jitter (max)	-.07**	.06**	0.006	0.03	-0.012	.06**	0.04
Jitter (mean)	-.08***	0.04	0.03	0.04	0.03	.09***	.07***
Shimmer (max)	-0.04	-.05*	-0.03	-0.03	-.05*	-0.03	-.07**
Shimmer (mean)	-.10***	-0.005	-0.02	0.01	0.004	.09***	0.02

Table 4: Overall correlation between speech features and PSS and FFMQ
*Note: FFMQ=Five Facet Mindfulness Questionnaire; O=Observing; D=Describing; AA=Acting with Awareness; NJ=Non-judging; NR=Non-reacting; PSS=Perceived Stress Scale. *p<.05;**p<.01;***p<.001*

Additionally, scores of PSS and PID subscales were all positively significantly correlated with Negative Affect scale at *r*(84)=.576, *p*<.001; with Detachment scale at *r*(84)=.505, *p*<.001; with Psychoticism scale at *r*(84)=.157, *p*<.001; with Antagonism scale at *r*(84)= .072, *p*=.001 and Disinhibition scale at *r*(84) = .404, *p*<.001. This shows that higher scores on PSS will result in higher scores on these subscales.

3.3. Linear Mixed Models

Linear mixed effects were examined with PSS FFMQ and PID as fixed predictors of speech features, with questionnaire scores nested within participants. Centered PSS scores entered as a fixed effect were able to significantly predict variance in mean F0, β=1.240534, *SE*=0.492199, *t*(84)=2.520, *p*=.014, *d*= 0.55. The results showed that higher PSS scores indicated increased mean F0 with a moderate effect size. Additionally, centered Acting with awareness scores marginally significant when predicting F0 mean, β=1.03758, *SE*=0. 545468, *t*(84)=1.902, *p*=.061, *d*=0.41. As such, participants were scored higher on Acting with awareness had lower F0 mean, with a moderately effect size. Moreover, only F0 mean and shimmer mean were significantly predicted by the PID subscales. Centered Negative affect scores significantly predicted F0 mean with a small effect size at β=3.27707, *SE*=1.27150, *t*(84)=2.577, *p*=.012, *d*=0.27. Centered Antagonism scores was marginally significant in predicting shimmer mean with a small effect size, β=.00264, *SE*=0.00125, *t*(84)=2.100, *p*=.039, *d*=0.22.

Speech Feature	**PID**				
	NA	DE	PSY	ANT	DIS
Intensity (max)	.069**	-.041	-.056*	.005	-.006
Intensity (mean)	.076***	-.077***	-.038	.087***	-.003
F0 (max)	.033	-.024	.78***	.009	.004
F0 (mean)	.193***	.077***	.080***	.055*	.119**
Jitter (max)	-.045*	-.083***	.006	.020	-.002
Jitter (mean)	-.031	-.103***	.012	.071**	.061**
Shimmer (max)	-.069**	-.010	.002	-.013	.019
Shimmer (mean)	-.060**	-.107***	.59***	.116***	.114***

Table 5: Overall correlations of speech features with PID
*Note: PID=The Personality Inventory of DSM-5; NA=Negative Affectivity; DE=Detachment; PSY=Psychoticisim; ANT=Antagonism; DIS=Disinhibition. *p<.05;**p<.01;***p<.001*

All other identified significant correlations between PSS and FFMQ with acoustic features were found to be insignificant in the mixed model analysis.

3.4. Age and Gender Effects
Both age and gender were significantly correlated with some of the speech features. Age significantly correlated with maximum F0 at $r(86)=.10$, $p<.001$, maximum jitter at $r(84)=.08$, $p<.001$ and maximum shimmer at $r(84)=.08$, $p<.001$. Thus, the variance in these three speech samples among participants can be explained by age difference.
Differences in gender also accounted for some of the variance in speech features, and was significant for maximum intensity at $t(84)=4.98$, $p<.001$, $d=1.09$ with females having higher intensity ($M=0.05$, $SD=0.05$) higher than males ($M=0.04$, $SD=0.03$). Mean intensity was also seen to be higher for females ($M=0.0053$, $SD=0.005$) than males ($M=0.0046$, $SD=0.004$) at $t(84)=3.17$, $p=.002$, $d=0.69$. Similarly, mean F0 was also higher for females ($M=50.66$, $SD=32.27$) than males ($M=29.58$, $SD=51.77$) at $t(84)=11.08$, $p<.001$, $d=2.42$. Lastly, both maximum and mean shimmer was greater for males than females at $t(84)=9.76$, $p<.001$, $d=2.13$ and $t(84)=13.01$, $p<.001$, $d=2.84$ respectively. Maximum shimmer scores for males were $M=0.93$, $SD=0.11$ and females was $M=0.87$, $SD=0.15$. The mean shimmer scores for males was $M=0.12$, $SD=0.03$ and females was $M=0.14$, $SD=0.04$.

4. Discussion

4.1. Summary of Findings
The study aimed to investigate the relationship between perceived stress, mindfulness, and personality with non-linguistic features of speech. Based on previous studies, four acoustic parameters were identified namely F0, intensity, jitter and shimmer, and they were found to be significant predictors of stress in speech. First, the association between perceived stress and acoustic features were investigated. The results demonstrated that higher perceived stress scores were positively correlated with mean F0. Second, the relationship between FFMQ and speech features were explored and the study found that higher scores on Acting with awareness subscale predicted lower scores on mean F0. Third, correlations between personality traits and speech features were examined to find that higher Negative affectivity subscale significantly showed an increase in mean F0 and higher Antagonism scores predicted higher shimmer mean. Although initially several other correlations between speech features and the self-reports were found, further analysis on them failed to achieve significance. This study will discuss the findings by looking at both significant predictors of speech features as well as significant correlations between self-reports and acoustic features (despite not being significant after linear mixed effects) to gain a comprehensive understanding of stress in speech.

4.2. General Discussion
Firstly, the study hypothesized that higher perceived stress scores would indicate higher F0 and intensity and lower jitter (H1). The findings are partly consistent as higher PSS scores predicted higher mean F0 with a moderate effect size ($d=0.55$). This association is in tandem with previous research that has documented F0 as a robust marker of stress in speech (Brenner & Shipp, 1988). F0 has shown to be predictive of stress in a variety of simulated stress environments and is now able to do so in natural setting as well. The study also found a negative correlation between perceived stress and jitter however, further analysis on it did not showcase jitter as a significant parameter to measure stress. Some of the previous studies found a decrease in jitter while others found an inconclusive association (Mendoza & Carballo, 1998); Even after this study, a significant relation between the two has not been established suggesting that jitter it is not the most effective measure of stress. Surprisingly, intensity was not found to associate with perceived stress. This is contrary to findings of (Hecker et al., 1968) that documented an increase in the parameter during stress induced conditions through cognitive workload and qualitative assessment of stress. As a result, only F0 is seen as a reliable predictor raising questions on the effectiveness of measuring stress through speech.
Secondly, the study hypothesized that the FFMQ facets would correlate with the acoustic features (H2). Higher Observe and Describe subscales were to predict increased stress via its speech features (increase in F0 and intensity; decrease in jitter) while Acting with awareness, Non-judging and Non-reacting subscales were to predict decrease through the same. The current findings found that only Acting with awareness subscale predicted stress through speech, hence only one facet supports the hypothesis. The results found that higher scores on subscale would significantly predicted lower mean F0 in speech which in turn indicated reduced stress levels. A plausible explanation is that individuals who score high on this subscale believe they have the abilities to undertake steps to combat stressful situations. This is further supported by a large correlation between Acting with awareness and PSS ($r=-.50$).
Contrastingly, the other four facets could not significantly predict any of the speech features. Observe subscale was negatively correlated with both mean intensity and mean F0 (opposite to what was expected) while Describe subscale failed to significantly correlate with any of the acoustic features. Since both these subscales are related to increasing sensitivity towards the perception of stress, it is strange that they were negatively correlated with PSS as well. On the other hand, higher Non-judging and Non-reacting scores indicated lower maximum intensity and increase in jitter respectively. These associations are in the expected direction, hence it is peculiar that they could not significantly predict any of the speech features despite having large correlations with PSS (Non-judging at $r=-.42$; Non-reacting at $r=.-54$). It can be suggested that high scores on these do not reduce the stress levels within individuals (gaged through objective measures) but helps them accept and appraise the situations calmly. This is consistent with Shapiro and Walsh's (2003) explanation that mindfulness practices assist in controlling and decreasing automatic responses but do not make one resilient towards stress.
Thirdly, it was hypothesized that shimmer would not be predicted by perceived stress or mindfulness facet scores (H3). Neither PSS nor FFMQ were able to predict shimmer as a valid marker of stress hence, the findings are support the hypothesis. However, significant correlations were found between shimmer and the two self-reports. Higher scores on PSS showed a decrease in shimmer indicating that shimmer indicates lower stress. The study also found that high scores on Non-reacting are positively associated with

shimmer, which is in sync with current finding since Non-reacting scores are negatively correlated with stress levels. However, the validity of shimmer as a parameter of measuring stress is yet inconclusive, consistent with past literature.

Lastly in an exploratory way, the study predicted that individual differences in personality would explain the variance in stress-related speech features. The results showed that Negative affectivity significantly predicted an increase in mean F0. The findings for Negative affectivity are in line with previous research that show that this dimension forms central part of "distressed" individuals who prone to react to situations with increased cortisol levels (Sher, 2004). A positive correlation with mean intensity further highlights that high scores on this subscale will demonstrate increased stress levels within individuals. A large positive correlation with perceived stress (r= .58) concludes that Negative affectivity is an important predictor of stress. The study also found Antagonism to significantly predict an increase in shimmer mean. However, in relation to stress, very little research has been conducted on Antagonism and its sub-parts namely Attention Seeking, Grandiosity, Deceitfulness and Manipulativeness. Additionally, because the findings of shimmer are also vague, it is unclear whether a rise in shimmer, predicted by Antagonism, is indicative of high or low stress. Moreover, the subscale was found to positively correlate with mean intensity and achieved a small correlation with perceived stress (r= .07).

Following which, Detachment subscale significantly correlated with all four speech features, higher scores on the dimension found lower scores for mean intensity, maximum and mean jitter and shimmer mean, increased scores in mean F0. Because increased F0 and reduced jitter showcase stress, it is plausible to imply that high scores on detachment result in greater stress levels. A large positive correlation with perceived stress (r= .51) supports this view. Additionally, high scores on Psychoticism obtained high scores in maximum and mean F0 and shimmer mean. This result is consistent with past literature as individuals with Psychoticism are found to have increased health-related risks such as smoking which eventually relate to stress in one's life. However, a precise relationship between Psychoticism and stress has not been found. Lastly, Disinhibition was found to positively correlate increase in shimmer mean and had a large correlation with perceived stress (r= .40). Once again, because the finding of shimmer is not comprehensive, it is unclear whether it causes an increase or decrease in stress. However, previous findings of the dimension have found that its sub-parts namely Risk taking and Impulsivity encourage high risk and sensation-seeking behaviour eventually attribute to stress within individuals.

Factors such as age and gender found to have an effect on the acoustic features. Age was correlated with maximums of F0, jitter and shimmer implying that as we grow older an increase in these acoustic features are observed. Similarly, females were found to have higher maximum and mean intensities as well as mean F0 while men were found to have higher maximum and mean shimmer. This finding has important implication since higher intensity and F0 contribute to increases in stress. Research by National Women's Health Resource Center (2003) showed that more than 90% of the women report moderate and higher levels of stress in their lives. As a result, this finding brought to light an alarming statistic that could now be investigated to improve well-being in society. For example, researchers could explore physiological and environmental factors that result in this discrepancy in speech features to further explain the differences in stress levels.

The current findings regarding the non-linguistic speech features suggest that whilst listeners can understand whether the speaker is stressed, F0 is the only acoustic parameter that is a reliable marker of stress. Although perceived stress correlated with the speech features in the expected direction, it was not able to predict the features. This implies that more focus should be given on why these features were not able to achieve significance. Moreover, researchers could incorporate the use of stress biomarkers such as hair and salivary cortisol in conjunction with speech recordings to have a better estimate of the stress experienced by individuals. Also, while conducting an experiment on stress, it is crucial to ensure that all participants indeed experienced stress and they were of similar levels, to avoid large inter-speaker variability (Hollien, 1990). This would involve monitoring of stress which is rather difficult unless physiological measures are used. Additionally, some researchers have not found HPA axis to reflect subjective response of psychological stress (Kirschbaum et al., 1996) doubting the role of PSS as a valid measure of stress. Researchers could use Trier Social Stress Test (TSST; Kirschbaum, Pirke & Hellhammer, 1993) that induces psychological stress as well as measures physiological reactions to stress that could effective in managing stress.

4.3. Limitations

Since the participants were required to record their speech three times a day over 10 days, the methodology appeared to be tiresome involving high level of commitment. The dropout rate of the participants was considerably high with 126 exclusions from the sample. This negatively affected the power of the analysis that was conducted. It is quite possible that because of this that self-report measures were not able to significantly predict acoustic features despite achieving significant correlations. Moreover, pre-processing of speech samples found several recordings to be blank or of poor quality, leading to further exclusions of data. Studies have shown that measurements of jitter could be adversely be affected by poor quality recordings while the distance between the microphone and speaker should also be at a fixed distance to avoid errors in intensity measurements (Kirchhuebel et al., 2011). Thus, a variety of parameters may have confounded the data yielding errors in recordings.

Another concern was attributed to false positives in results considering the large size of the data (1970 recordings) and various measures used to test the recordings. In order for a measure to predict an acoustic feature in a robust manner, all its functionals should have altered in the expected direction. However, in most cases either the mean or the maximum of a particular acoustic changed. Despite the Bonferroni corrections, the tests may still result in a Type II error. The study uses a novel measuring technique and explores multiple factors, hence, the findings need to be replicated to before drawing inferences from the results. Furthermore, 56% of the participants were non-native English speakers, which could be a cause of concern because the participants were asked to speak aloud sentences in English. Reading aloud sentences in a language that is not most commonly used can be inferred as cognitive overload, which inherently could be a causing psychological stress. Hence, having a non-native English sample can be problematic.

4.4. Future Research

Although various acoustic features have been studied, there are other non-linguistic features that can be investigated. These include quality of voice, pause time between syllables and reduced speech can also provide information about an individual's stress level (Buchanan et al., 2014). One of the reasons for pauses between syllables is for lexical retrieval (Levelt, 1983) and stress induced cortisol is found to disrupt memory retrieval (Wolf, 2009), hence increase in pauses may indicate higher stress levels, thus the relationship between these can be looked into. Less attention has been given on quality of voice which plays an important role in F0 and intensity, thus can be explored.

The initial move towards the development of complementary stress-reducing interventive steps lies in the ability to accessibly measure stress through speech, though it is unable to identify the underlying stressor. Mindfulness-based strategies such as meditation can be integrated into the smartphone application, so long as these findings are ascertained as robust. Through a self-monitoring mechanism, ones' awareness of stress levels can be improved upon. Before clinical consultations, one can employ the speech-based stress appraisal systems in order to help recognize the markers of chronic stress by reporting when the stress-related symptoms begin to occur. The speech-based stress appraisal is virtually free compared to other physiological measurements, and perhaps more unbiased than self-report questionnaires. Thus it can widely be used to measure stress levels of various people in a go. Hence, it addresses the imminent problem of increasing stress levels in society.

4.5. Conclusion

The study of stress through speech is an area of research that has received relatively less attention, with physiological measures of stress such as cortisol levels in hair and saliva receiving more. However, measuring these is not convenient since one cannot predict exactly when stressful situations will occur. Hence, the novel technique allowing the assessment of stress through a smartphone application makes it low-cost and accessible, ensuring the measurement of stress levels of various people at a single time. This study asked participants to record their speech over 10 days and these speech features were investigated using perceived stress, mindfulness and personality dimensions. This study proposes that a combination of these speech features predict stress levels within individuals. However, the validity of speech under stress is not fully established and more research in this field will advance the extent to which speech features measure stress**.**

5. References

i. Andreou, E., Alexopoulos, E. C., Lionis, C., Varvogli, L., Gnardellis, C., Chrousos, G. P., et al. (2011). Perceived stress scale: Reliability and validity study in Greece. International Journal of Environmental Research and Public Health, 8, 3287-3298.

ii. Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. Assessment, 13(1), 27–45.

iii. Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. (2008). Construct validity of the Five Facet Mindfulness Questionnaire in meditating and nonmeditating samples. Assessment, 15(3), 329-342.

iv. Baken, R. J., &Orlikoff, R. F. (2000). Clinical measurement of speech and voice. San Diego, CA: Singular Publishing Group.

v. Bandura, A. (1992). Exercise of personal agency through the self-efficacy mechanism. In R. Schwarzer (Ed.), Self-efficacy. Thought Control of Action (pp. 3-38). Washington: Hemisphere.

vi. Booth-Kewley, S. & Vickers, R. R. (1994). Associations between major domains of personality and health behavior. Journal of Personality, 62, 281-298.

vii. Bränström, R., Duncan, L. G., & Moskowitz, J. T. (2011). The association between dispositional mindfulness, psychological well-being, and perceived health in a Swedish population-based sample. British Journal of Health Psychology, 16(2), 300-316.

viii. Brenner, M. & Shipp, T. (1988). Vocal stress analysis. Mental-State Estimation, 1987, 363–376. NASA Conference Publication 2504.

ix. Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. Journal of Personality and Social Psychology, 84(4), 822–848. doi:10.1037/0022-3514.84.4.822.

x. Cannon, W. B. (1915) Bodily Changes in Pain, Hunger, Fear and Rage: an account of recent researches into the function of emotional excitement. New York: Appleton.

xi. Carmody, J., & Baer, R. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. Journal of Behavioral Medicine, 31(1), 23–33.

xii. Cohen S., Kessler R. C., & Gordon, U. L. (1995). Strategies for measuring stress in studies of psychiatric and physical disorder. In S. Cohen, R. C. Kessler, & U. L. Gordon (Eds.), Measuring Stress: A Guide for Health and Social Scientists (p. 3-26). New York, NY: Oxford University Press.

xiii. Cohen, S., Kamarck, T., &Mermelstein, R. (1983). A global measure of perceived stress. Journal of Health and Social Behavior, 24(4), 385-396.

xiv. Collins English dictionary. Accessed 12 April 2010 from http://www.collinslanguage.com/

xv. Costa, P. T., Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO-Five-Factor Inventory (NEO-FFI) Professional Mannual. Odessa, FL: Psychological Assessment Resources.

xvi. Dumitru, V. M., &Cozman, D. (2010). The relationship between stress and personality factors, 4(1), 34–39.

xvii. Eyben, F., Weninger, F., Gross, F., & Schuller, B. (2013, October). Recent developments in openSMILE, the Munich open-source multimedia feature extractor. In Proceedings of the 21st ACM International Conference on Multimedia (p. 835-838). New York, NY: Association for Computing Machinery.

xviii. Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X. & Zhang, K. (2014). Effect of stress and personality on dangerous driving behaviour among Chinese drivers. Accident Analysis and Prevention, 73(1), 34-40.

xix. George, J.M. (1992). The role of personality in organizational life: Issues and evidence. J.Manag.,18:185-213.

xx. Harrington, J., Palethorpe, S., & Watson, C. I. (2007). Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. INTERSPEECH, 2753- 2756.

xxi. Headley, B. & Wearing, A. (1989). Personality, life events, and subjective well-being: toward a dynamic equilibrium model. Journal of Clinical Psychology, 52(4), 383-387.

xxii. Hecker, M. H., Stevens, K. N., von Bismarck, G., & Williams, C. E. (1968). Manifestations of task-induced stress in the acoustic speech signal. The Journal of the Acoustical Society of America, 44, 993–1001.

xxiii. Hicks, J. W. Jr. (1979). An acoustical/temporal analysis of emotional stress in speech (Unpublished doctoral dissertation). University of Florida, Florida, United States.

xxiv. Huttunen, K., Keränen, H., Väyrynen, E., Pääkkönen, R., &Leino, T. (2011). Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. Applied Ergonomics, 42(2), 348–57.

xxv. Jessen, M. (2006). Einfluss von Stress auf Sprache und Stimme: UnterbesondererBerücksichtigungpolizeidienstlicherAnforderungen. Idstein, Germany: Schulz-Kirchner Verlag GmbH.

xxvi. Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. General Hospital Psychiatry, 4, 33–47.

xxvii. Kirchhuebel, C., Howard, D. M., &Stedmon, A. W. (2011). Acoustic correlates of speech when under stress: Research, methods and future directions. International Journal of Speech, Language and the Law, 18(1), 75-98.

xxviii. Kirkcaldy, B. D., Shephard, R. J., &Furnham, A. F. (2002). The influence of Type A behavior and locus of control upon job satisfaction and occupational health. Personality and Individual Differences, 33, 1361–1371

xxix. Kirschbaum, C., Wolf, O., &Hellhammer, D. (1998). Adrenocortical responsiveness to psychosocial stress in humans: Sources of interindividual differences. In D. S. Krantz, A. Baum (Eds.), Technology and Methods in Behavioral Medicine (p. 29–45). New Jersey, NJ: Erlbaum.

xxx. Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., &Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychological Medicine, 42, 1879-1890.

xxxi. Kudielka, B. M., &Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. Biological psychology, 69(1), 113-132.

xxxii. Kurniawan, H., Maslov, A. V., &Pechenizkiy, M. (2013). Stress detection from speech and galvanic skin response signals. Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems, 209–214.

xxxiii. Lively, S. E., Pisoni, D. B., Van Summers, W., &Bernacki, R. H. (1993). Effects of cognitive workload on speech production: Acoustic analyses and perceptual consequences. The Journal of the Acoustical Society of America, 93(5), 2962-2973.

xxxiv. Matthews, G., Saklofske, D. H., Costa, P. T., Deary, I. J. &Zeidner, M. (1998). Dimensional models of personality: A framework for systematic clinical assessment. European Journal of Psychological Assesment, 14(1), 36-49.

xxxv. Mendoza, E., & Carballo, G. (1998). Acoustic analysis of induced vocal stress by means of cognitive workload tasks. Journal of Voice, 12(3), 263–273.

xxxvi. Ortleb, R. (1937). An objective study of emphasis in oral reading of emotional and Unemotional Material, Speech Mono, 4, 56-58.

xxxvii. Protopapas, A. and Lieberman, P. (1997) Fundamental frequency of phonation and perceived emotional stress. Journal of the Acoustical Society of America 101: 2267–2277.

xxxviii. Russell, E., Koren, G., Rieder, M., & Van Uum, S. (2012). Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. Psychoneuroendocrinology, 37(5), 589-601.

xxxix. Sauve, B., Koren, G., Walsh, R., Tokmakejian, S., &Uum, S. H. M. V. (2007). Measurement of cortisol in human hair as a biomarker of systematic exposure, 30(5), 183-192.

xl. Scherer, K.R. (1979) Nonlinguistic vocal indicators of emotion and psychopathology. In C.E. Izard (ed.) Emotions in Personality and Psychopathology 495–529. New York: Plenum Press.

xli. Scherer, K.R. (1986) Vocal a ect expression: a review and a model for future research. Psychological Bulletin 99(2): 143–165.

xlii. Tiffin, J. and Steer, M. D. (1937). An Experimental Analysis of Emphasis, Speech Mono . , 4,69-74.

xliii. Trancoso, I, & Moore, R. (1995), Proceedings of the ESCA-NATO Tutorial and Research Workshop on Speech under Stress. Lisbon, Portugal: ESCA.

xliv. Yang, Y., Koh, D., Ng, V., Lee, F. C., Chan, G., Dong, F., & Chia, S.E. (2001). Salivary cortisol levels and work-related stress among emergency department nurses. Journal of Occupational and Environmental Medicine, 43(12), 1011-1018.

xlv. Yao, X., Jitsuhiro, T., Miyajima, C., Kitaoka, N. & Takeda, K. (2015). Modeling of physical characteristics of speech under stress. IEEE Signal Processing Letter, 22(10), 1801-1805.

Annexure

Appendix A

Translated from German, originally in Burkhardt et al.'s (2005) database of emotional speech.

1	One plus one is two.
2	Six plus eleven is seventeen.
3	DNA is a molecule that encodes the genetic instructions used in the development and functioning of all known living organisms.
4	Attempts to bring elements of natural language grammar into computer programming have produced new programming languages.
5	Learning also encompasses many other techniques that seek to summarize and explain key features of the data.
6	She was tall and thin, with brown hair.
7	Mathematics is the study of topics such as numbers, structure, space, and change. There is a range of views among mathematicians and philosophers as to the exact scope and definition of mathematics.
8	The universe cannot be read until we have learned the language and become familiar with the characters in which it is written.
9	Mathematics is used throughout the world as an essential tool in many fields, including natural science, engineering, medicine, finance and the social sciences.
10	The evolution of mathematics might be seen as an ever-increasing series of abstractions, or alternatively an expansion of subject matter.
11	Everyone knows that 5 plus 5 is 10.
12	The tablecloth is lying on the fridge.
13	She will hand it in on Wednesday.
14	Tonight I could tell him about my idea.
15	The black sheet of paper is located up there besides the piece of timber.
16	What about the bags standing there under the table?
17	I will just discard this and then go for a drink with them.
18	It will be in the place where we always store it.
19	We keep a comprehensive range of replacement spare parts for up to ten years.
20	Their training gives them expertise and knowledge of our appliances that only a manufacturer can provide**.**

Stimuli list of neutral English sentences used for speech recordings.

Appendix B

Formulae and calculation of speech features (Compiled by P. Schneider, SOMA Analytics).

1. Source-Filter Model

Human speech consists of two components: the linguistic unit of words, and the voice itself. While words carry most of the content, the voice transports more subtle information. The production of speech starts when air is released into the vocal cords from the lungs. Through the contractions and expansions of the vocal cords, the inflowing air gets excited. Following this, the excited air flows through the vocal tract and is shaped through amplifying certain harmonics. In signal processing, the excitation of air in the vocal cords is called source and the reshaping of the airflow in the vocal tract, a filter. Hence, this model is often referred to as the Source-Filter model.

2. Low-level Features

In order to compute low-level features, a given speech signal is broken down into frames of length 25ms - 32ms. All the low-level features are then computed over each frame. These are further explained below:

i. Fundamental Frequency. From the source-filter model we know that the air is excited by the vibrations of the vocal cords. Fundamental frequency measures the periodicity of the vibrations.

ii. Energy. Energy in speech can be thought of as energy referred to in physics. It is computed over a small time frame as the sum of the squared signal:

$$E = \sum_{n=t}^{T} x^2[n]$$

Voiced sounds tend to have more energy than unvoiced ones since more energy is needed to force the air stream through the vocal cords.

iii. Jitter. This describes the deviation in period length of the fundamental frequency between frames. Note that a frequency can be represented as number of cycles in a given time period. Thus the period length is simply the amount of time one cycle needs. For example, a frequency of 10Hz relates to 10 cycles per second, which is a period length of 100 ms. The formula for Jitter is given as follows:

$$J = \frac{1}{N-1}\sum_{i=1} |T_i - t_{i+1}|$$

Where T refers to the period length of the fundamental frequency, and N refers to the number of frames.

iv. Shimmer. Shimmer represents the frame to frame variability of the fundamental frequency amplitudes. Amplitudes of a spectrum give more information about the contribution of a certain frequency to the entire signal. The unit for amplitude is decibel, which is computed in the term $20log$(A). The formula for shimmer can be defined as:

$$S = \frac{1}{N-1}\sum_{i=1} \left| 20\log\left(\frac{A_{i+1}^{F_0}}{A_i^{F_0}}\right) \right|$$

v. Further Processing. All the features measured on a frame basis are smoothed by applying a simple moving average over three consecutive values.

3. High-level Features

High-level features were computed over the entire sequence of frames that make up an entire recording. In the given data set, the minima, maxima and the first four centered moments were computed for each low-level feature over all frames.