An examination into the effects of speech rate on perceived stress in monolingual and bilingual populations

Rachel Francis

Supervised by: Astrid Schepman

April 2019

Abstract

Stress is the body's response to adverse or demanding circumstances and can cause physical changes such as increased respiratory rate and increased vocal cord muscle tension which can affect speech production and the acoustic properties of speech such as speech rate (duration). Acoustic properties such as duration and intensity act as cues in stress judgements, with duration proving to be the factor that provides the greatest fluctuation in these judgements. In the present study, 55 monolingual and bilingual (Spanish and English speaking) participants listened to 6 audio files spoken at 3 varying speeds in both English and Spanish and rated how stressed they perceived the speaker to be. Contrary to what was predicted, the results demonstrated high intra-cultural similarities in terms of perceptions of stress. As hypothesised, higher stress ratings were attributed to the faster spoken files, although they were also attributed to those files spoken in Spanish. There were interactions between the speed of audio and participant group, language spoken and participant group, speed of audio and language spoken and finally speed of audio, language being spoken and participant group. These results demonstrate that speech rate has significant effects on perceptions of stress and also suggest that the effect of speech rate on these perceptions varies between languages However, previous literature would suggest that the acoustic properties of speech are affected differently in real life scenarios compared to when speech is manipulated artificially. Suggesting that further research should endeavour to avoid electronically manipulated audios, instead capturing naturally occurring audio files.

KEY WORDS:

SPEECH RATE BILINGUIALISM STRESS STRESS JUDGEMENT SPEECH MANIPULATION

Introduction

On average people speak 16,000 words a day (Huynh, 2014), expressing many emotions and attitudes through this speech (Mozziconacci & Hermes, 1999). This research will examine perceptions of stress based solely on voice, specifically speech rate, and analyse the effects of speech rate on these perceptions in both monolingual and bilingual populations.

Stress

Stress has been given many contrasting definitions by various different people ranging from medical professionals, social scientists and psychologists (Selye, 1987). According to the English Oxford Living Dictionary stress can be defined as "A state of mental or emotional strain or tension resulting from adverse or demanding circumstances". At the most basic level, stress is the way that our bodies respond to pressures from an adverse situation, life event or circumstance and what causes this response can vary drastically between individuals according to numerous factors such as social and economic circumstances, living environment and genetic makeup (Mental Health Foundation, 2018). In order for an event to be labelled a stressor it must be perceived as one by the individual however, this perception differs amongst individuals (Selye, 1987). Although, some common features do exist between individuals in terms of triggers of stress including new or unexpected experiences, something threatening a feeling of self and a feeling of lack of control over a given situation (Mental Health Foundation, 2018).

There are many varying forms of stress ranging from cognitive stress, emotional stress, workload task induced stress and situational stress (Bliss, 1956; Epel, Daubenmier, Moskowitz, Folkman & Blackburn, 2009; Hansen, 1996; Matthews & Campbell, 1998; Sorg & Whitney, 1992). It is the existence of these various kinds of stress and the lack of clear operational distinctions between them that raises issues in studying the effects of stress on speech (Scherer, Grandjean, Johnstone, Klasmeyer & Bänziger, 2002). Not only does this insufficient specification of the type of stress being studied lead to issues in making generalisations and drawing conclusions amongst studies, but many research projects also demonstrate a strong neglect of accounting for individual differences in terms of susceptibility to stress (Scherer et al, 2002). Due to the many studies that fail to acknowledge this it is difficult to uncover research within this field that provides reliable results that can be applied to a wider population (Giddens, Barron, Byrd-Craven, Clark & Winter, 2013).

Furthermore, the difficulty in terms of inducing stress in a reliable and ethical manner in a laboratory setting causes additional issues, with it being highlighted that the stress induced in a laboratory environment differs greatly to that of a real stressor (Ruiz, Absil, Gramatica, Harmegnies, Legros & Poch, 1995). Studies using experimentally induced stress generally produce weaker results, possibly due to the fact that laboratory induced stress is much less powerful than naturally

occurring stress (Tolkmitt & Scherer, 1986), therefore complicating research and reducing reliability and generalisability (Ruiz et al, 1995).

Many studies have delved into the effects of stress on speech through the analysis of audio recordings. Streeter, Macdonald, Apple, Krauss and Galotti (1983) investigated the effects of stress in this manner through the study of the telephone conversations of two Consolidated Edison employees during New York's 1977 blackout. The results of this study proved to be inconclusive as is most research within the field. Whilst stress appeared to affect the intensity, speech rate and fundamental frequency of one employee's voice, it had no such effect on the other.

The effects of stress on speech can also be examined through the audio analysis of aeroplane or helicopter pilots' discussions in dangerous or life-threatening situations (Šulc & Remek, 1986; Dewaele & Furnham, 2000). However, it is difficult to identify trends or form solid conclusions based on this research due to the minimal amount of data that there is to be studied. Although, some researchers have aimed to overcome this through the study of speech as a result of simulated situations similar to those stated above (Jones Jr, 1990; Huttunen, Keränen, Väyrynen, Pääkkönen & Leino, 2011) and despite the lack of real life evidence, suggest that through a combination of these simulated situations and real life audios, we can analyse the way in which speech is affected by stress and the differences caused by laboratory induced stress and real life stressors.

The background literature therefore would appear to suggest that analysing the effects of stress on speech is more ethically sound through the analysis of audios recorded during a naturally occurring stressful situation, as opposed to inducing stress in participants and examining the effects of this. This research aims to follow this example by measuring participants perceptions of stress levels based on speech rate, rather than having their own speech rate recorded whilst being put under stress.

Vocal Properties

Normal speech can be referred to as speech produced in a quiet environment with no distractions or task obligations (Nwe, Foo & De Silva, 2003). On the other hand, stress in speech arises as a result of various contributing factors such as fatigue, environmental noise or distractions and fluctuations in an individual's emotional state. These stressors can cause physical changes such as an increase in respiratory rate, irregular breathing and increased vocal cord muscle tension, all of which can affect speech production and the acoustic properties of speech (Nwe, Foo & De Silva, 2003). Previous research confirms this, concluding that the presence of stress changes phoneme production with regards to glottal source factors, pitch, intensity, duration and spectral shape (Hansen, 1988).

Years of research highlight the numerous varying aspects of speech such as speech rate, fundamental frequency, fluency, pronunciation, pitch and duration

(Bachorowski & Owren, 1995; Solé, 2007; Steeneken & Hansen, 1999; Wu, Parsons & Narayanan, 2010), that serve as non-verbal signalling systems and allow us to draw conclusions on an individual, their personal attributes and the various emotions that they may be experiencing (Allport & Cantril, 1934; Apple, Streeter & Krauss, 1979; Bou-Ghazale & Hansen, 1994; Collins & Missing, 2003; Hansen, 1996; Oksenberg, Coleman & Cannell, 1986; Rothkrantz, Wiggers, Van Wees & Van Vark, 2004). Stress alters the functioning of the speech production process which results in "irregular vocal fold movement, vocal tract articulator perturbation, variation in airflow from the lungs, changes in subglottal air pressure, and other vocal systems modifications." (Bou-Ghazale & Hansen, 1996). Consequently, affecting the vocal properties of an individual's voice.

Many researchers have focused on how stress affects these properties and from this research it can be seen that various forms of stress, including some of those previously mentioned such as situational and emotional stress, impact speech production and cause significant variations in speech in comparison to neutral speech (Hansen, 1996). However, some aspects of speech vary differently according to the type of stress being experienced; for example, the effects on the pitch of somebody's voice differ when an individual is experiencing emotional stress compared to when they experience workload task stress (Godin & Hansen, 2008). This is generally due to the speaker attempting to counteract the effects of workload stress on their speech, much unlike the case in many instances of speech produced under emotional stress. Under emotional stress the speaker may experience unconscious changes in the acoustic properties of their speech or may even consciously modify their vocal qualities, in order to communicate information about their emotional state.

The most frequently investigated acoustic parameter is fundamental frequency, which aids in conveying information about an individual's state (Tolkmitt & Scherer, 1986). Numerous studies analysing the effects of naturally stressful situations on fundamental frequency indicate that mean fundamental frequency increases under extreme stress (Demenko, 2008; Sigmund, 2013; Streeter et al, 1983) and although less conclusive, studies examining the effects of stress induced in a laboratory setting produce similar results, demonstrating a strong tendency in the same direction (Protopapas & Lieberman, 1997; Tolkmitt & Scherer, 1986).

With regards to previous research focused on the effects of pitch on perceptions of personal attributions, specifically truthfulness, it has been discovered that people view lower pitched speakers as being more truthful and credible than higher pitched speakers (Streeter, Krauss, Geller, Olson & Apple, 1977; Apple, Streeter & Krauss, 1979). Apple, Streeter and Krauss (1979) suggest that the link found between pitch and truthfulness is possibly a result of a listener's perception of higher pitch being an indication of stress. Consequently, they attribute this stress to attempted deception: deeming the higher pitched speaker less truthful than the lower pitched speaker.

Apple et al. (1979) also focused on the effects of stress on speech rate and how this affected truthfulness ratings: concluding that faster paced speakers were viewed as being more truthful than slower paced speakers. However, from these findings, it appears that the effect of speech rate on truthfulness judgements is not linear, as is that of pitch, but rather it has an inverted U-function in that the slower and faster speech were perceived as less credible and truthful than the 'normal' paced speech (with the slower speech being perceived as the least credible). This pattern has also been highlighted in research conducted by Smith, Brown, Strong and Rencher (1975).

Buchanan, Laures-Gore and Duff (2014) studied the effects of stress on speech properties and highlighted that speech fluency (as measured by the number and duration of pauses per minute) is significantly reduced when under stress. Despite these findings concluding that stress reduces the number and duration of pauses, further research suggests that it is in fact associated with an increased use of filled pauses: um, er, uh (Christenfeld & Creager, 1996), highlighting the inconsistencies within the results of studies in this area of research.

It is evident from these previous studies that vocal properties are in some way affected by stress and that this stress causes physical changes within the speech production process. However, considering the existing literature, it is also clear that most stress studies are limited in scope, often only studying one or two participants and analysing the same number of speech parameters; pitch and fundamental frequency being more often than not the favoured vocal properties studied (Hansen, 1996; Tolkmitt & Scherer, 1986). Within the field it is not unusual for conflicting results to be produced due to the many differences between studies experimental designs, levels of actual or simulated stress or researcher's interpretations of results (Hansen, 1996). As a result of this conflict, extensive research is required in order to increase the number of findings that can be analysed and compared: Consequently, minimising the existing gap and increasing overall knowledge of the effects of stress on the acoustic properties of speech.

Speech Rate

Whilst the findings of some studies appear to suggest that speech rate increases as a result of stress (Hansen, 1989; Hollien et al, 1993 as cited in Kirchhübel, Howard & Stedmon, 2011; Siegman, 1993 as cited in Kirchhübel, et al, 2011), other researchers have suggested that the speed of speech is not affected by stress (Scherer, Grandjean, Johnstone, Klasmeyer, & Bänziger, 2002). However, it appears to be that when data has proven insignificant, negative conclusions have been drawn. Therefore, it is not the case that no difference has been found with regards to speech rate when participants are stressed versus when they are not, but more so that the results produced did not prove to be significant enough to conclude that stress does significantly affect speech rate.

Despite the fact that numerous studies will have negated the existence of this effect due to insignificant results, they cannot deny that a change does exist in speech rate in the presence of stress. An interesting question here is whether or not this change necessarily needs to be proven to be significant in a laboratory setting for it to exist in real life scenarios, or for it to have an effect on how people perceive a speaker's speech rate and whether or not they deem it an indication of stress.

Previous studies analysing speech under stress have failed to consider the varied ways in which the duration of speech can be affected. There is a lack of consideration for the numerous aspects of a sentence and also individual words that can be affected should speech rate speed up or slow down: For example, whether or not it is the duration of whole words or individual phoneme-classes (i.e. vowels or consonants) that is altered (Hansen, 1996). Kuwabara (1997) reported that in comparison to consonants, the duration of vowels is reduced significantly more so from slow to fast speech. Thus, it is possible that rather than an overall increase or decrease in mean speech duration acting as an aid in drawing conclusions on an individual's stress levels, it is in fact a variation in one of these specific aspects of speech.

It has also been shown that speech rate can dramatically impact the degree of variation in pronunciation (Fosler-Lussier & Morgan, 1998; Greenberg & Fosler-Lussier, 2000), as well as affecting the presence of deletions, insertions, and coarticulation effects (Benzeghiba et al, 2007). For example, deletions are increased in fast speech whilst insertion errors are increased in slow speech (Martinez, Tapis, Alvarez & Leon, 1997; Nanjo & Kawahara, 2004). It is important to consider here that as is the case with the above-mentioned phoneme-class differences, this effect of speech rate on pronunciation, deletions, insertions and substitution errors could be the contributing factor that alters people's perceptions of an individual's stress levels; not the overall mean duration of their speech.

Stress as a subset of emotion

Considering that the main body of research in this field is based on emotion perception, it is vital to consider the relationship between stress and emotion, whether the two concepts are inclusive of each other or whether stress forms part of the category of emotion. Whilst a number of theorists have proposed the existence of a group of basic emotions, the exact number of these is a controversial topic, as is the mere concept of their existence. Ekman, Friesen and Ellsworth (2013) conducted a review of decades of research and concluded that the majority of findings suggest the existence of six basic emotions. However, other research has suggested that there are seven (Ekman & Keltner, 1970) and that any other emotion' is a combination of the basic emotions and can be categorised as non-basic (Ekman, 1992).

The majority of the literature does in fact relate to perceptions of emotions. However, despite the fact that stress does not appear within any of the basic

emotion groups aforementioned, some psychologists have argued that stress should be considered part of the larger topic of emotion (Lazarus, 1993), as the concept of emotion includes that of stress. Although it is clear that the topic of emotions is much broader than that of stress, with stress being more limited in breadth and depth.

All things being considered, it therefore seems reasonable to apply any previously identified trends and findings produced from research into the effects of stress on vocal properties, and the effects of these qualities on perceptions of emotions, to the study of stress.

Emotion perceptions based on vocal properties

It has been suggested that certain vocal qualities can help form the basis of judgements on demographic characteristics, with those such as gender and age being accurately identified from the voice (Giles, Scherer & Taylor, 1979; Pear, 1931). These demographic characteristics also appear to influence personal attribute judgements although it has been suggested that the effect of vocal qualities on personal attributions may be dependent on the sex of the speaker (Addington 1968; as cited in Oksenberg et al 1986).

Cahn (1990) studied perceptions of emotion through the manipulation of synthesized speech parameters in order to convey various emotions. The study focused on whether the speaking style, created through the manipulation of parameters such as loudness, speech rate, pitch range and breathiness amongst others, affected how well somebody could judge the emotion being portrayed. The findings suggest that people are able to determine some emotions more consistently and easily than others based on speaking style. For example, sadness was reported as being the most consistently recognised emotion by participants. Other emotions included anger, surprise and disgust.

With this being said, it has been suggested that speech rate, variations in pitch and acoustic intensity are the most influential vocal properties on perceptions and judgements of a speaker, their emotions and their personal attributes (Fry, 1958; Scherer, 1972A; Scherer, 1972B). With regards to speech rate it has been observed that faster spoken speech leads people to attribute highly potent emotions to the speaker, such as anger and happiness whilst slower tempo speech leads to attributions of sadness, disgust and boredom (Pavlenko, 2007).

Further psychologists have also proposed that a faster speech rate is linked to personal characteristics such as persuasiveness (Mehrabian & Williams, 1969; Packwood, 1974) and increased intelligence and knowledge (Miller, Maruyama, Beaber & Valone, 1976). Moreover, Fry (1958) further studied perceptions of stress in speech and highlights that vocal properties such as duration and intensity also act as cues in stress judgements, with duration proving to be the factor that provides the greatest fluctuation in these judgements.

The research within this area is very limited and as a result of this, a huge gap exists with regards to perceptions of stress levels based on speech rate. This is due to the fact that the majority of research focuses on perceptions of basic emotions and mainly studies the effects of stress on other acoustic properties such as fundamental frequency and pitch. The results of the current study will help to overcome this existing lack of research, possibly encouraging further study into the perceptions of stress levels based on vocal qualities, specifically speech rate.

Cross-cultural differences

Very little is known about the cross-cultural differences in vocal emotion processing and the effect that vocal property manipulation has on this process due to the lack of research within the area. Whilst these cross-cultural differences in emotion processing and perception judgement have previously been studied, the focus of the research has mainly been on the perception of emotion through facial expressions (Ekman, 1993; Ekman & Friesen, 1971; Massaro & Ellison, 1996; Masuda et al, 2008; Sauter, Eisner, Ekman & Scott, 2010) and the research into these differences between the Spanish and English cultures is almost non-existent. Although, some psychologists have investigated these judgements cross-culturally between speakers of languages such as German, American, Polish and Japanese (Beier & Zautra, 1972; Breitenstein, Lancker & Daum, 2001; Scherer, 1972B). On the whole, the data from these studies highlight overall higher intra-cultural similarities compared to cross-cultural differences in emotional ratings given and perceptions of emotions based on voice, suggesting that despite language differences, an innate understanding of emotion judgements exists across cultures and languages.

It could be hypothesised that the current study should produce results similar to those of previous cross-cultural research focused on emotion should stress be considered a subset of emotion as aforementioned. In that, more intra-cultural similarities will be found than will cross-cultural differences, with regards to the ratings given and participants' perceptions of stress levels.

On the other hand, findings from a study conducted by Hirschberg, Biadsy, Rosenberg, Carlson & Strangert (2008) suggest that although there are some acoustic-prosodic correlates which are common across cultures, there are other acoustic-prosodic and lexical correlates which are specific to languages; finding that in some cases, correlations of judgements for speech in participants' native language are quite different from the correlations of the same participants rating a foreign language. In this case this was with regards to American's ratings of Arabic speech, a language which they had no knowledge of. Similarly, in the current study, monolingual participants will have no knowledge of the second language being used: Spanish. Despite the fact that this study focused on perceptions of charisma rather than stress, the current study could be expected to produce similar results, with monolingual participants perceiving the audios

spoken in a foreign tongue as less stressed than those spoken in their native language due to it being unfamiliar to them.

Considering the overall lack of research within this area, the conflicting results and the inadequacy of recent data, regarding not only the effects of stress on speech rate and perceptions of stress based on speech rate, but also the differences in these effects and perceptions when considering cross cultural populations, the current study aims to provide findings that can begin to fill the gap that currently exists within the field and encourage further research. Particularly with regards to differences in the perceptions and judgements of English and Spanish participants.

Study Rationale

As discussed above, there is currently a very limited amount of literature focused specifically on the effects of stress on speech rate, perceptions of stress based on speech rate and studies into the differences of this in Bilingual and Monolingual populations is almost non-existent.

The present study will look into perceptions of stress based on speech rate and also analyse whether any differences exist with regards to these perceptions in both monolingual and bilingual participants who can speak Spanish and English. This study will also attempt to determine whether or not speech rate is of greater significance in English than it is in Spanish for bilingual participants when judging a speaker's stress levels.

Hypotheses

I hypothesise that (1) Monolingual participants will rate the individual speaking in the audio as being more stressed during the faster speaking files regardless of the language spoken, (2) Bilingual participants will rate the individual speaking in the audio as being more stressed during the faster speaking files but will also rate the speaker as being more stressed in the English files than in the Spanish files; regardless of their native language. I hypothesise this difference between languages based on previous findings that when spoken, Spanish has a significantly increased rate of syllables per second compared to English (Pellegrino, Coupé & Marsico, 2011). Thus suggesting that bilingual participants, who are familiar with both languages, will be accustomed to Spanish being a faster paced language; therefore, they may be less likely to attribute higher stress levels to the speaker when speaking faster in Spanish due to the fact that they already view Spanish as being a faster paced language than English, consequently not noticing the increases in speed as easily as they do in English.

Method

Participants

Based on previous research in this area and also taking into account the possibility that there may not be a large number of Native Spanish bilingual participants readily available, the aim was to recruit 20 participants for each of the three groups in order to make a total of 60 participants. Participants consisted of 55 University of Chester students: 18 monolinguals, 20 Native English Bilinguals and 17 Native Spanish Bilinguals.

This study adhered to BPS ethical guidelines and received ethical approval from the Department of Psychology Ethics Committee at the University of Chester. Monolingual participants were recruited via opportunity sampling through the University's Research Participation System (RPS) and were awarded 2 RPS points upon completion of the study. In addition, in order to recruit additional monolingual participants, students within the researcher's psychology lectures were also recruited verbally by the researcher. Native English Bilingual Participants were recruited by the researcher who verbally asked the students if they would participate during a lecture as they were studying the same course. Native Spanish Bilingual Participants were initially informed of the study via an email sent by the administrator of the University's Language Department. Those students who wished to participate then got in touch with the researcher via email or face to face during lectures in order to arrange a suitable date and time for participation.

Materials and Stimuli

The audio stimuli used in this research were created by the researcher; three texts were written in English and three in Spanish and matched in terms of the general topic areas and neutral content as to avoid upsetting or distressing participants. Upon completion of the texts, the researcher then recorded them as she speaks both languages. Once all of the texts had been recorded, the speed of four of the audio files was manipulated through the use of the computer software Praat (Boersma & Weenink, 2019). Of the three audios in each language, one was slowed down and one was sped up, with the remaining audio in each language being unedited. The duration of the files was edited using the duration manipulation within the software. An unedited audio file has a duration of 1: in order to speed up two of the audio files (one in each language), the duration was manipulated to 0.93, whilst the duration of those that were slowed down was manipulated to 1.07. The overall length of the audio files ranged from 16 seconds to 32 seconds, with an average length of 26.2 seconds.

The online questionnaire was created through the use of Online Surveys (https://www.onlinesurveys.ac.uk) and distributed in the lab: it was displayed on the individual computers to which each participant was assigned, and they completed the questionnaire in the laboratory. The researcher created a questionnaire which measured participants' views on the perceived stress levels of the speaker. The questionnaire consisted of six pages, one for each audio file. On each of the six

pages, there were nine statements which could be rated from 1-9 on a Likert scale. Each item consisted of two words, one at each end of the scale (E.G the number one would indicate “calm” and number 9 for the same item would indicate “tense”). Participants were required to provide a number rating based on how closely they thought the statements reflected the speakers voice in the respective audio file: For example, if a participant thought that the speaker sounded extremely tense, they would provide a rating of 9. The nine items consisted of five stress measuring items (various synonyms of the word stressed such as “tense” and “on edge”) and four distractor items such as “old”, “happy”, “shy”.

Procedure

The experiment was conducted in a quiet lab room which contained six computers. Each participant was assigned to a computer and dividers prevented them from being able to see the computer screen of the participants either side of them. On average, there were four participants for each data collection session, with five being the maximum number allowed to participate in any given session.

The experiment began with the researcher providing information sheets to all participants which described the purpose of the study and the format it would follow. Alongside the information sheet, participants were also provided with a consent form which they were required to sign before partaking in the study in order to gain informed consent. The information sheet also gave information on how the data would be used and informed the participants that they had the option to refuse participation should they no longer wish to take part.

Participants then listened to the 6 audio files which were played aloud, via a speaker, by the researcher who ensured that the volume was at a comfortable level. The audios were played in an order that had been predetermined by the researcher, in order to ensure that the responses provided by participants corresponded to the same audio file for each data collection session. Participants were required to provide ratings for each file via the online questionnaire. Each file was rated immediately after being played. Upon completion of the ratings, participants were given a debrief sheet, explaining the true purpose of the study and informing participants that the speed manipulation of the audio stimuli had been withheld from them prior to and during the study. The debrief form highlighted that participants should refer themselves to student wellbeing in the unlikely event that the study affected them in any way. It also provided points of contact for both the researcher and the research supervisor should the participant feel that they needed to discuss the study further.

Ethical Considerations

In accordance with BPS guidelines, it is vital to supply participants with as much information as possible prior to the study (British Psychological Society [BPS], 2014). However, deception is sometimes required when a key element of the research would be compromised should participants be fully aware of the aim of the study:

although it is important that, should deception be utilised, the study is designed to protect the dignity and autonomy of participants (BPS, 2014). Whilst this study did involve deception, the information being withheld from participants would not have caused any psychological stress. Regardless of this, in order to adhere to BPS guidelines, once the data had been collected, all participants were provided with an appropriate form of debriefing which explained why the deception was necessary, the true purpose of the study and provided contact details of the researchers should they feel the need to contact them following the study.

Design and analysis

A 2x3x3 mixed design was employed for the study. The between subjects independent variable was the participant group (with three levels; monolinguals, native English bilinguals and native Spanish bilinguals), whilst the within subjects independent variables were the language being spoken on the audio (with two levels; English or Spanish) and the speed of the audio file (with three levels; normal, slow or fast). The dependent variable was the ratings provided by participants via the online questionnaire.

The responses provided by participants were exported as a spreadsheet into Excel and the distractor items were omitted as the ratings provided for these factors were not relevant to the study. The remaining 5 items consisted of stress ratings and the scores provided by each participant over these 5 items were averaged in order to generate 1 stress rating per audio file per participant. This data was then transferred to SPSS v24 and analysed through a 2x3x3 mixed ANOVA to determine whether any interactions existed. Post-hoc follow up paired sample t-tests were then conducted in order to test each hypothesis.

Results

Data Preparation

Once all data had been collected and exported into Microsoft Excel, the 4 distractor items were omitted as they were not relevant to the research question. The 5 stress ratings given by each participant were averaged to provide 1 stress rating per participant per audio file: No items were reverse coded, allowing for means to be easily calculated. No data was deemed an outlier and subsequently no data points were excluded. Missing data for any participant were substituted with the mean of the remaining stress ratings provided by that participant. The refined data was then copied into SPSS and analysed.

Descriptive Statistics

Below is a table demonstrating the mean ratings given by Monolingual participants, Bilingual Native English and Bilingual Native Spanish participants for the 3 varied

speeds of audio files in the 2 chosen languages. A higher score indicates that a higher level of stress has been perceived.

Table 1: Demonstrates the means (SDs) for average stress rating given by participants based on Language being spoken and Speed of speech (minimum score=1, maximum score=9).

	Participant Group	Mean
English Normal	Monolingual	3.07 (0.86)
	Bilingual Native English	2.92 (1.00)
	Bilingual Native Spanish	3.25 (1.12)
	Total	**3.07 (0.99)**
Spanish Normal	Monolingual	4.84 (0.99)
	Bilingual Native English	3.40 (1.56)
	Bilingual Native Spanish	3.12 (0.91)
	Total	**4.00 (1.36)**
English Slow	Monolingual	2.33 (1.13)
	Bilingual Native English	3.06 (1.21)
	Bilingual Native Spanish	2.24 (0.93)
	Total	**2.57 (1.15)**
Spanish Slow	Monolingual	3.24 (1.14)
	Bilingual Native English	2.93 (1.08)
	Bilingual Native Spanish	2.65 (1.10)
	Total	**2.95 (1.12)**
English Fast	Monolingual	3.30 (1.11)
	Bilingual Native English	3.41 (1.48)
	Bilingual Native Spanish	4.45 (2.04)
	Total	**3.69 (1.63)**
Spanish Fast	Monolingual	4.52 (2.07)
	Bilingual Native English	2.74 (1.10)
	Bilingual Native Spanish	3.98 (1.34)
	Total	**3.71 (1.70)**

Main Effects and Interactions

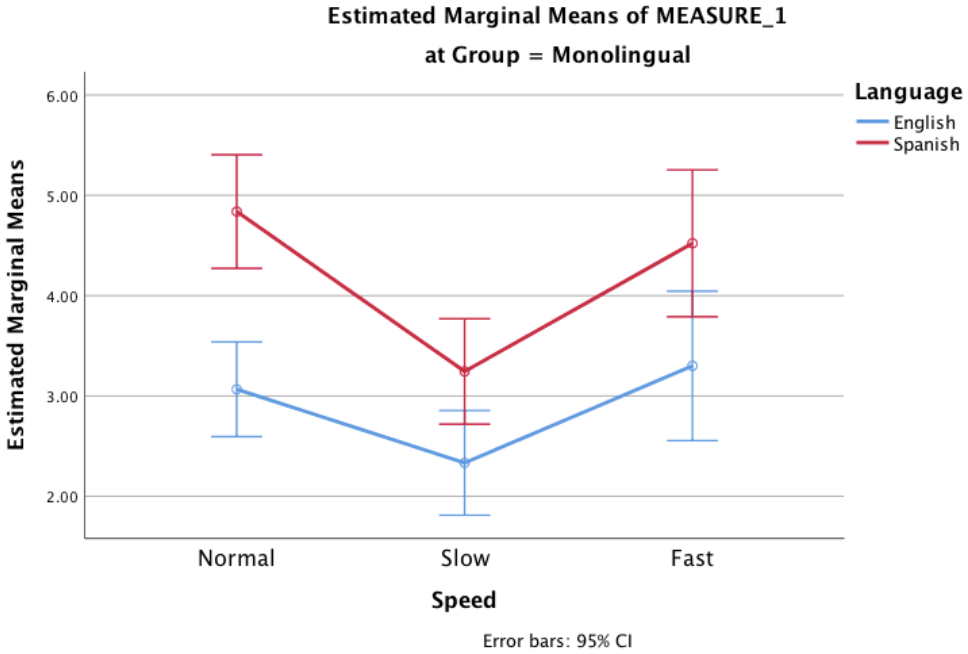
A 2x3x3 mixed ANOVA was conducted to analyse the differences in the average stress ratings given by participants for the 6 audio files and revealed a significant main effect of speed ($F(2,104) = 23.33$, $p < 0.001$, $\eta_p^2 = .31$) in that a higher stress rating was attributed to the faster spoken audio files. Additionally, there was a significant main effect of language ($F(1,52) = 11.15$, $p = 0.002$, $\eta_p^2 = .18$) in that higher stress ratings were provided for the Spanish spoken audio files.

There was an interaction between the speed of the audio and participant group ($F(4,104) = 6.64$, $p < .001$, $\eta_p^2 = .20$). In addition, there was an interaction between the language spoken and participant group ($F(2,52) = 10.43$, $p < .001$, $\eta_p^2 = .29$). There was also an interaction between the speed of the audio and the language being spoken ($F(2,104) = 8.40$, $p < .001$, $\eta_p^2 = .14$). Finally, there was an interaction between the speed of the audio, the language being spoken and the participant group ($F(4,104) = 4.32$, $p < .003$, $\eta_p^2 = .14$)

Monolingual Participants

Figure 1 demonstrates that Monolingual participants perceived the speaker as more stressed during the Spanish spoken audio files than in the English files regardless of them being spoken at the same rate. With regards to the English files their perceptions of the speaker's stress levels increased with the increasing rate of speech. However, when considering the Spanish files, the trend was not similar to that of the English files: although participants perceived the slower speaking file as being the least stressed of the 3 Spanish files, with the normal and fast paced files being viewed as more stressed than this, the normal paced speech was perceived as being more stressed than the faster speech.

Figure 1: The interaction between speed of audio and language spoken for monolingual participants.

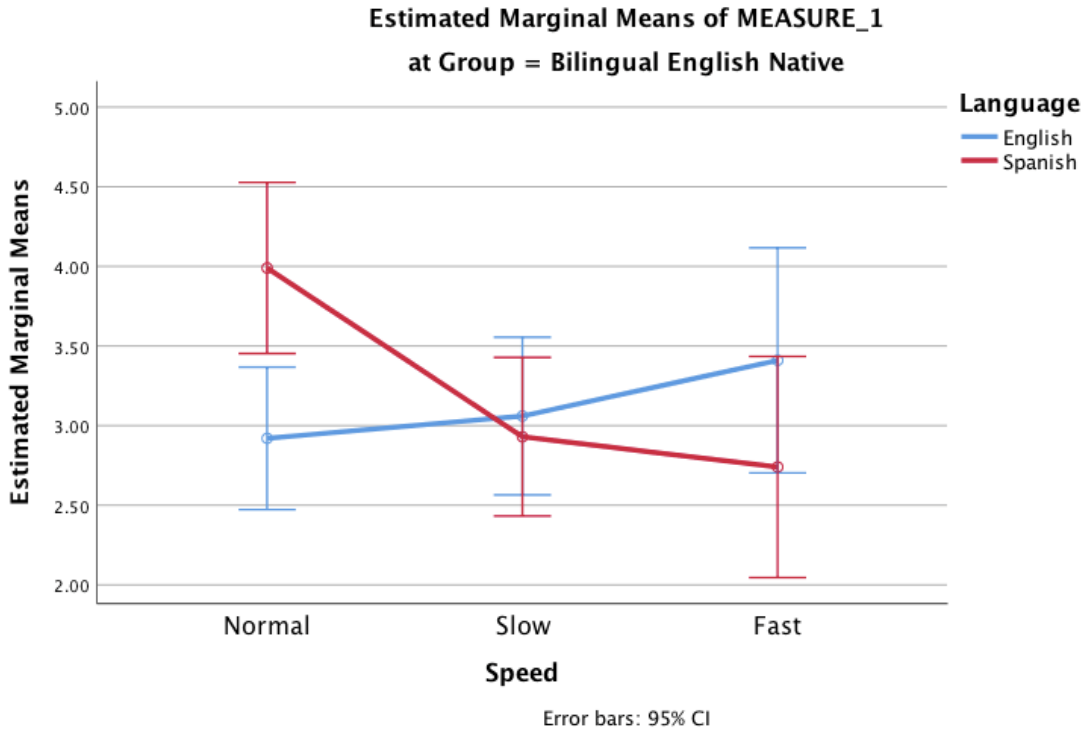


Bonferroni corrections were applied in order to reduce the chances of type I error as multiple pair wise tests were performed on a single set of data. The Bonferroni correction divided the otherwise used alpha level of 0.05 by 18, consequently reducing alpha to 0.003. Post hoc paired samples *t*-tests confirmed that the only ratings given by monolingual participants that were significantly different, following the Bonferroni corrections, were those given for the Spanish files spoken at the normal and slower speeds. With the normal speed file being rated as more stressed than the slower spoken file ($t(17) = 1$, $p<0.001$)

Bilingual English Native Participants

Figure 2 demonstrates a crossover with regards to the stress ratings provided by bilingual English native participants for each of the two languages. Regarding the English audios, **figure 1** highlights an increasing trend in stress ratings given with the normal speech rate receiving the lowest rating and the faster spoken file receiving the highest. However, the trend reverses when considering the Spanish language files: with the highest stress levels being attributed to the normal paced speech and the lowest stress levels being attributed to the faster paced speech.

Figure 2: The interaction between speed of audio and language spoken for bilingual English native participants.

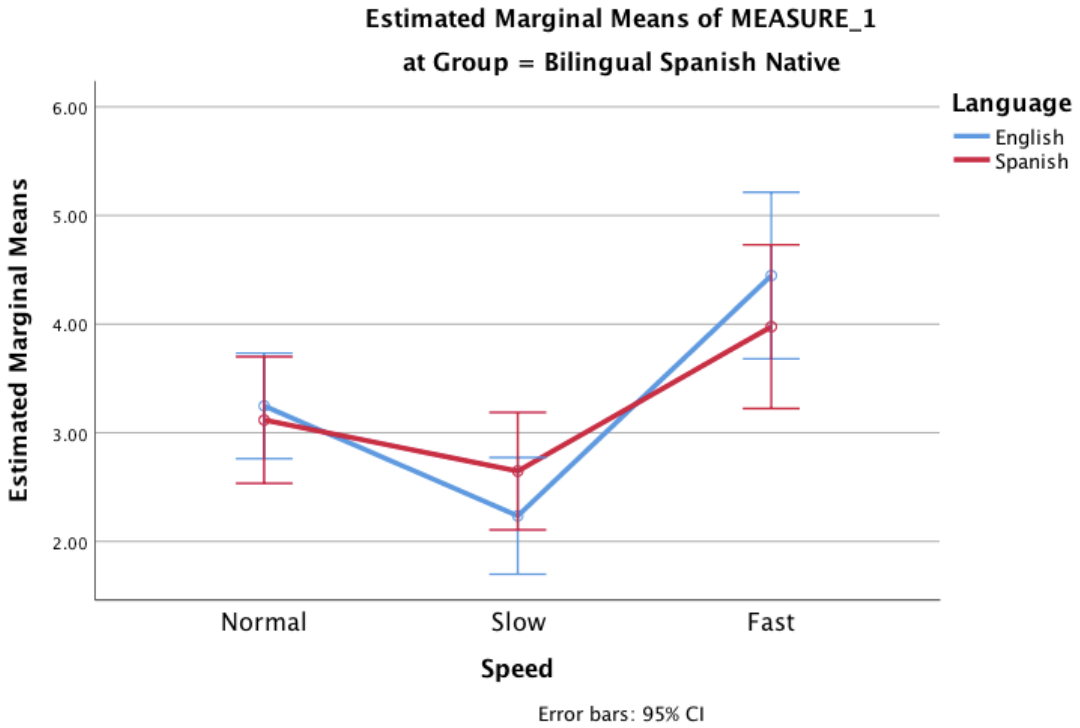


Post hoc paired samples *t*-tests confirmed that the ratings given by bilingual English native participants were significantly different for both the normal vs slow spoken Spanish file and the normal vs fast spoken Spanish file. The t-test confirmed that stress ratings were significantly higher for the normal spoken file than the slower spoken file ($t(19) = 3.74$, $p=0.001$), as well as unexpectedly being significantly higher for the normal spoken file than the faster spoken file ($t(19) = 3.60$, $p=0.002$).

Bilingual Spanish Native Participants

Figure 3 demonstrates that Bilingual Spanish native participants believed the speaker to be more stressed during the faster spoken files whilst the slower spoken files were deemed the least stressed; although the language perceived as the most stressed at each speech rate varied.

Figure 3: The interaction between speed of audio and language spoken for bilingual Spanish native participants.



Post hoc paired samples *t*-tests confirmed that the ratings given by bilingual Spanish native participants were significantly different for the normal vs slow, normal vs fast and slow vs fast audio files spoken in English. The t-test confirmed that stress ratings were significantly higher for the normal spoken file than the slower spoken file ($t(16) = 4.75$, $p<0.001$), significantly higher for the faster spoken file than the normal spoken file ($t(16) = -3.62$, $p=0.002$) and significantly higher for the faster spoken file compared to the slower spoken file ($t(16) = -5.01$, $p<0.001$).

Discussion

The results of the current study did not show a significant main effect of participant group on stress ratings given. However, they did highlight a significant main effect of both speed and language: with the speaker being perceived as more stressed when speaking Spanish and when speaking at an increased speech rate. Not only were significant main effects found for both speed and language, but significant interactions were also identified amongst all factors: with each factor interacting with the two other factors and all 3 factors interacting concurrently.

These results provide support for the hypotheses to some extent. With regards to the first hypothesis that Monolingual participants would rate the speaker as being more stressed during the faster spoken audio files, the results for the English spoken audio files support this theory as the speaker was viewed as more stressed as the speech rate was increased. However, unexpectedly, the same trend was not identified for the Spanish spoken files: although the file spoken at the slowest rate was given the

lowest stress rating, the faster file was rated as less stressed than the normal paced audio. Therefore, these findings do not entirely support the hypothesis as it suggested that the language spoken would have no effect on the ratings provided by monolingual participants; however, the speaker was viewed as having significantly higher stress levels when speaking Spanish at each varying speed.

With regards to the second hypothesis and firstly addressing the native Spanish bilinguals, the results partly support the predictions that bilingual participants, regardless of their native language, would rate the speaker as more stressed during the faster spoken files and more so when speaking English compared to when speaking Spanish. Although the responses given by Spanish native bilingual participants resulted in the predicted trend, with stress ratings being the lowest for the slowest spoken files and highest for the fastest spoken files, the predicted interaction between language and speed of audio file was not supported as speakers stress levels were not always judged as being higher for the files spoken in English. For the audio files that were slowed down, the speaker was perceived to be less stressed when speaking Spanish than when speaking English. However, despite this, considering the mean stress ratings given, this population seem to be more susceptible to noticing differences in speech rate in their non-native language as predicted.

Finally, the results again partly support the second hypothesis with regards to the native English bilinguals: stress ratings were higher for the fastest spoken English file, although contradictory to predictions the speaker was rated as less stressed during the normal paced file than during the slower spoken file. In addition, when considering the Spanish spoken files, although the slower spoken file was perceived as less stressed than the normal paced file, the faster spoken audio received the lowest stress rating of all 3. Moreover, although it was hypothesised that the interaction between speed and language would be stronger for the English files than the Spanish, it was in fact the opposite trend that was identified in the files in which the speeds had been manipulated. The only audio files which provide support for the hypothesis were those in which the speech rate was not electronically modified.

As the data in this field of research is very limited, it is difficult to see whether or not the results fall in line with previous work as it is greatly lacking. However, the current findings do appear to agree with the concept previously suggested by Fry (1958) that, speech rate has a significant influence on somebody's ability to judge another individual's emotions. Not only this, but the idea that more highly potent emotions are attributed to faster speech (Pavlenko, 2007) is supported by the findings that, with regards to the English spoken files, on average the highest stress levels were awarded to the fastest spoken file. Although this is not the case with the Spanish files.

Previous studies imply that higher intra-cultural similarities compared to cross cultural differences exist in emotion judgements, although it should be noted that these conclusions have been drawn based on the judging of facial expressions in addition to voice (Beier & Zautra, 1972; Breitenstein, Lancker & Daum, 2001; Scherer, 1972B). The current findings appear to agree with past research in that, overall, more

similarities exist between groups than do differences. Although only the Spanish native bilingual participants consistently attributed higher stress levels to the increasing speech rates in both languages.

These results contradict the hypothesis that there would be significant cross-cultural differences between monolingual and bilingual participants. However, one explanation for why these cross-cultural differences were not found may be explained through the findings of a study conducted by Benzeghiba et al (2007) which reported on a series of experiments that investigated how speakers produce fast speech and how listeners perceive said speech. They summarised that the main question, with regards to this difference, is how perceptions differ when considering naturally produced fast speech compared to artificially time compressed speech.

It has been shown that speech rate can dramatically impact the degree of variation in pronunciation (Fosler-Lussier & Morgan, 1998; Greenberg & Fosler-Lussier, 2000), as well as affecting the presence of deletions, insertions, and coarticulation effects (Benzeghiba et al, 2007). For example, deletions are increased in fast speech whilst insertion errors are increased in slow speech (Martinez, Tapis, Alvarez & Leon, 1997; Nanjo & Kawahara, 2004). It is important to consider here that this effect of speech rate on pronunciation, deletions and insertion and substitution errors could be the contributing factor that alters people's perceptions of an individual's stress levels; not the overall mean duration of their speech. If this were to be the case, with regards to the current study, the changes in speed of speech may not have the same effect on participants ratings as it would do in a real life scenario, due to this factor being manipulated electronically rather than the audio being naturally spoken at a faster rate: consequently, having no effect on the parameters of deletion and pronunciation (etc.).

In addition to this, some researchers believe that speech fluency (as measured by the number and duration of pauses per minute) is reduced when under stress (Buchanan, Laures-Gore & Duff, 2014), whilst others believe that stress in fact causes an increase in the use of filled pauses (Christenfeld & Creager, 1996). Despite these conflicting opinions, the fact that these differences in speech fluency have been identified when individuals are under stress again highlights the difficulties faced in the current study by electronically manipulating speech rate in order to convey stress. For example, although the software speeds up or slows down the overall duration of speech, it does not add in any filled pauses such as "ums" or "uhs", nor does it lengthen specific pauses made by the speaker: the duration as a whole is manipulated.

Furthermore, in many studies analysing speech under stress, there is a lack of consideration for the numerous aspects of a sentence and also individual words that can be affected should speech rate speed up or slow down: For example, whether or not it is the duration of whole words or individual phoneme-classes (i.e. vowels or consonants) that is altered (Hansen, 1996). Kuwabara (1997) reported that in comparison to consonants, the duration of vowels is reduced significantly more so from slow to fast speech. Thus, it is possible that rather than an overall increase or decrease in mean speech duration acting as an aid in drawing conclusions on an

individual's stress levels, it is in fact a variation in one of these specific aspects of speech. Therefore, the manipulation of the speed of audio files in the current study may not have achieved the desired effect as the mean speech duration has little effect on participants' perceptions. If specific phoneme-classes of the files were to be manipulated, the results may have differed greatly. Reiterating the aforementioned theory that artificially manipulating the audio files in the current study could prevent the naturally occurring differences that are apparent in faster spoken speech, resulting in no effect on participants stress judgements.

Practical applications

These findings present practical implications and applications with regards to public speaking as it would appear to be that in English, higher stress levels are attributed to faster spoken speech. With regards to public speakers such as politicians or campaign managers and leaders, this information on how others perceive stress based on speech rate is of great benefit as it will give a greater and more in depth understanding of how their speech rate affects the effect that they have on an audience's perceptions of them and what it is that they discuss. The idea of stress contagion suggests that it is possible for an individual who is stressed to "pass on" their stress to observers through their speech (Buchanan, Bagley, Stansfield & Preston, 2012), therefore suggesting that should a public speaker speak at a faster pace, viewers will perceive them to be stressed about either the situation itself or what it is that they are talking about. If this were to be within a political environment it may lead the listeners to believe that the politician is distressed about the subject that they are talking about, consequently causing the listener to become stressed about the subject. Findings such as those presented in the current study allow public speakers to evaluate the ways in which listeners view speech rate and how significant this is in influencing their perceptions of stress: enabling them to work on establishing an optimum speaking rate in order to avoid this stress contagion.

Limitations

The findings of the current study are limited due to the fact that speech rate was electronically manipulated; leading to the above-mentioned issues of a clear disregard of the changes in speech rate that naturally occur in faster spoken speech such as increased deletions and fluctuations in speech fluency. Because of this it is not possible to truly accept the results as conclusive, as it is not possible to attribute the changes in stress ratings solely to the varied speech rates. The use of audio files in which the speaker's speech rate is naturally at an increased or decreased rate may result in more accurate findings for future studies.

Furthermore, the sample size of the current study was relatively small, mainly due to the limited Spanish native bilingual individuals that were available, causing issues with the generalisability of the results. Further research should aim to recruit more participants in order to provide more accurate results. A recommended ratio of respondents to variables within factor analysis would be a minimum of 10:1, meaning a total of 90 participants, as a larger sample size will help diminish any error within

the data and tends to provide more precise, stable results (MacCallum, Widaman, Zhang & Hong, 1999; Yong & Pearce, 2013). Although it has also been recommended than the number of participants should be at least 100 (Kline, 2014).

Future Research

Future researchers within this field should endeavour to further explore the relationship between speech rate and perceptions of stress, analysing the differences and similarities between monolingual and bilingual populations. Specifically, studying this interaction when using natural speech that has not been electronically manipulated. This could be explored by capturing audios of an individual speaking when under various levels of naturally occurring stress (e.g. whilst giving a public speech compared to when talking with friends about a holiday), consequently resulting in different speech rates, but also resulting in differences in deletion errors, pronunciation (etc.). It would also be beneficial to recruit a larger sample in order to reduce the chance of any error and increase the accuracy of results allowing them to be applied to a wider population. It is clear to see that the findings from this study provide a foundation for future research that can be built upon by further studies.

Conclusion

In conclusion, the literature would suggest that the changes in acoustic properties and vocal qualities caused by artificial manipulation are not the same as those caused naturally through stress, therefore implying that the most effective way to study perceptions of stress based on qualities such as speech rate would be for participants to judge natural, unedited sound files. As previously mentioned, laboratory induced stress is much less powerful than naturally occurring stress (Tolkmitt & Scherer, 1986), providing further support for the study of naturally occurring stressed speech such as that of an aeroplane or helicopter pilots' discussions in a dangerous or life-threatening situation. Although the current study did not provide support for all of the proposed hypotheses, the results most definitely provide an intriguing insight into perceptions of stress based on speech rate, the differences in these perceptions when considering different languages and also how these perceptions differ in bilingual populations in their native and second language.

In summary, the present study adds recent data to the existing body of limited literature, identifying higher levels of intra-cultural similarities than cross-cultural differences in emotion perceptions. As research into these perceptions based on speech rate and differences in the perceptions of monolingual and bilingual populations is limited, the present study provides the foundation for a new area of research. However, due to the methodological limitations previously discussed, replication and adaptation is required in order to arrive at more conclusive results.

References

Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *The Journal of Social Psychology*, *5*(1), 37-55. https://doi.org/10.1080/00224545.1934.9921582

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, *37*(5), 715. http://dx.doi.org/10.1037/0022-3514.37.5.715

Bachorowski, J. A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological science*, *6*(4), 219-224. https://doi.org/10.1111/j.1467-9280.1995.tb00596.x

Beier, E. G., & Zautra, A. (1972). Identification of vocal communication of emotions across cultures. Retrieved from https://files.eric.ed.gov/fulltext/ED056604.pdf

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., ... & Rose, R. (2007). Automatic speech recognition and speech variability: A review. *Speech communication*, *49*(10-11), 763-786. https://doi.org/10.1016/j.specom.2007.02.006

Bliss, E. L., Migeon, C. J., BRANCH, H. C., & Samuels, L. T. (1956). Reaction of the adrenal cortex to emotional stress. *Psychosomatic Medicine*, *18*(1), 56-76. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.489.9995&rep=rep1&type=pdf

Boersma, P & Weenink, D (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.46, retrieved 3 January 2019 from http://www.praat.org/

Bou-Ghazale, S. E., & Hansen, J. H. (1994, April). Duration and spectral based stress token generation for HMM speech recognition under stress. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on* (Vol. 1, pp. I-413). IEEE. DOI: 10.1109/ICASSP.1994.389268

Bou-Ghazale, S. E., & Hansen, J. H. (1996). Generating stressed speech from neutral speech using a modified CELP vocoder. *Speech Communication*, *20*(1-2), 93-110. https://doi.org/10.1016/S0167-6393(96)00047-7

Breitenstein, C., Lancker, D. V., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition & Emotion*, *15*(1), 57-79. https://doi.org/10.1080/02699930126095

British Psychological Society. (2014). *Code of Human Research Ethics.* Retrieved from: https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/BPS%20Code%20of%20Human%20Research%20Ethics.pdf

Buchanan, T. W., Bagley, S. L., Stansfield, R. B., & Preston, S. D. (2012). The empathic, physiological resonance of stress. *Social neuroscience*, *7*(2), 191-201. https://doi.org/10.1080/17470919.2011.588723

Buchanan, T. W., Laures-Gore, J. S., & Duff, M. C. (2014). Acute stress reduces speech fluency. *Biological psychology*, *97*, 60-66. https://doi.org/10.1016/j.biopsycho.2014.02.005

Cahn, J. E. (1990). The generation of affect in synthesized speech. *Journal of the American Voice I/O Society*, *8*, 1-19. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.435.9599&rep=rep1&type=pdf

Christenfeld, N., & Creager, B. (1996). Anxiety, alcohol, aphasia, and ums. *Journal of Personality and Social Psychology*, *70*(3), 451. http://dx.doi.org/10.1037/0022-3514.70.3.451

Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal behaviour*, *65*(5), 997-1004. https://doi.org/10.1006/anbe.2003.2123

Demenko, G. (2008, May). Voice stress extraction. In *Proceedings of the Speech Prosody 2008 Conference*. Retrieved from https://pdfs.semanticscholar.org/9d56/57339e1aafb15c81036cfbab636bd8f449ff.pdf

Dewaele, J. M., & Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. *Personality and Individual differences*, *28*(2), 355-365. https://doi.org/10.1016/S0191-8869(99)00106-3

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200. https://doi.org/10.1080/02699939208411068

Ekman, P. (1993). Facial expression and emotion. *American psychologist*, *48*(4), 384. http://dx.doi.org/10.1037/0003-066X.48.4.384

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, *17*(2), 124. http://dx.doi.org/10.1037/h0030377

Ekman, P., Friesen, W. V., & Ellsworth, P. (2013). *Emotion in the human face: Guidelines for research and an integration of findings*. Elsevier. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=dOFFBQAAQBAJ&oi=fnd&pg=PP1&dq=Emotion+in+the+human+face:+Guidelines+for+research+and+an+integration+of+findings.+&ots=zhDGtNZBVb&sig=mOZrmkNaVzvef4ozvCU6PLKGqkk#v=onepage&q=Emotion%20in%20the%20human%20face%3A%20Guidelines%20for%20research%20and%20an%20integration%20of%20findings.&f=false

Ekman, P., & Keltner, D. (1970). Universal facial expressions of emotion. *California mental health research digest*, *8*(4), 151-158. Retrieved from

https://www.paulekman.com/wp-content/uploads/2013/07/Universal-Facial-Expressions-of-Emotions1.pdf

English Oxford Living Dictionaries. *Definition of Stress in English.* Retrieved December 4, 2018 from https://en.oxforddictionaries.com/definition/stress

Epel, E., Daubenmier, J., Moskowitz, J. T., Folkman, S., & Blackburn, E. (2009). Can meditation slow rate of cellular aging? Cognitive stress, mindfulness, and telomeres. *Annals of the new York Academy of Sciences*, *1172*(1), 34-53. https://doi.org/10.1111/j.1749-6632.2009.04414.x

Fosler-Lussier, E., & Morgan, N. (1998). Effects of speaking rate and word frequency on conversational pronunciations. In *Modeling Pronunciation Variation for Automatic Speech Recognition*. Retrieved from https://www.isca-speech.org/archive_open/archive_papers/mpv_98/mpv8_035.pdf
Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, *1*(2), 126-152.); https://doi.org/10.1177/002383095800100207

Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F., & Winter, A. S. (2013). Vocal indices of stress: a review. *Journal of Voice*, *27*(3), 390-e21. https://doi.org/10.1016/j.jvoice.2012.12.010

Giles, H., Scherer, K. R., & Taylor, D. M. (1979). Speech markers in social interaction. Retrieved from https://archive-ouverte.unige.ch/unige:102169/ATTACHMENT01

Godin, K. W., & Hansen, J. H. (2008). Analysis and perception of speech under physical task stress. In *Ninth Annual Conference of the International Speech Communication Association*. Retrieved from https://www.isca-speech.org/archive/archive_papers/interspeech_2008/i08_1674.pdf

Greenberg, S., & Fosler-Lussier, E. (2000, May). The uninvited guest: Information's role in guiding the production of spontaneous speech. In *Proceedings of the Crest Workshop on Models of Speech Production: Motor Planning and Articulatory Modelling* (pp. 129-132). Retrieved from http://http.icsi.berkeley.edu/ftp/pub/speech/papers/crest00-uninv.pdf

Hansen, J. H. (1996). Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition. *Speech communication*, *20*(1-2), 151-173. https://doi.org/10.1016/S0167-6393(96)00050-7

Hansen, J. H. (1989, March). Evaluation of acoustic correlates of speech under stress for robust speech recognition. In *Bioengineering Conference, 1989., Proceedings of the 1989 Fifteenth Annual Northeast* (pp. 31-32). IEEE. DOI: 10.1109/NEBC.1989.36683

Hirschberg, J. B., Biadsy, F., Rosenberg, A., Carlson, R., & Strangert, E. (2008). A cross-cultural comparison of American, Palestinian, and Swedish perception of charismatic speech. https://doi.org/10.7916/D8M3342B

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), 348-357. https://doi.org/10.1016/j.apergo.2010.08.005

Huynh, J. (2014). *Study Finds No Difference in the Amount Men and Women Talk.* Retrieved from https://ubrp.arizona.edu/study-finds-no-difference-in-the-amount-men-and-women-talk/

Jones Jr, W. A. (1990). *An evaluation of voice stress analysis techniques in a simulated AWACS environment* (No. AFIT/CI/CIA-90-115). AIR FORCE INST OF TECH WRIGHT-PATTERSONAFB OH SCHOOL OF ENGINEERING.

Kirchhübel, 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 & the Law*, *18*(1). Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a228727.pdf

Kline, P. (2014). *The New Psychometrics: Science, Psychology and Measurement*. Routledge. Retrived from https://books.google.co.uk/books?hl=en&lr=&id=9BisAgAAQBAJ&oi=fnd&pg=PR9&ots=OsM30Daer1&sig=n3DAbfU_aEGm-XJISYiyfLCTnBo#v=onepage&q&f=false

Kuwabara, H. (1997). Acoustic and perceptual properties of phonemes in continuous speech as a function of speaking rate. In *Fifth European Conference on Speech Communication and Technology*. Retrieved from https://www.isca-speech.org/archive/archive_papers/eurospeech_1997/e97_1003.pdf

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual review of psychology*, *44*(1), 1-22. Retrieved from https://www.annualreviews.org/doi/pdf/10.1146/annurev.ps.44.020193.000245

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, *4*(1), 84. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.536.6317&rep=rep1&type=pdf

Martinez, F., Tapias, D., Alvarez, J., & Leon, P. (1997). Characteristics of slow, average and fast speech and their effects in large vocabulary continuous speech recognition. In *Fifth European Conference on Speech Communication and Technology*. Retrieved from https://www.isca-speech.org/archive/archive_papers/eurospeech_1997/e97_0469.pdf

Massaro, D. W., & Ellison, J. W. (1996). Perceptual recognition of facial affect: Cross-cultural comparisons. *Memory & cognition*, *24*(6), 812-822. Retrieved from https://link.springer.com/content/pdf/10.3758/BF03201104.pdf

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: cultural differences in the perception of facial emotion. *Journal of personality and social psychology*, *94*(3), 365. http://dx.doi.org/10.1037/0022-3514.94.3.365

Matthews, G., & Campbell, S. E. (1998, October). Task-induced stress and individual differences in coping. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 11, pp. 821-825). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193129804201111

Mehrabian, A., & Williams, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *Journal of Personality and Social psychology*, *13*(1), 37. http://dx.doi.org/10.1037/h0027993

Mental Health Foundation. (2018). *Stress.* Retrieved from https://www.mentalhealth.org.uk/a-to-z/s/stress

Miller, N., Maruyama, G., Beaber, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of personality and social psychology*, *34*(4), 615. http://dx.doi.org/10.1037/0022-3514.34.4.615

Mozziconacci, S. J., & Hermes, D. J. (1999). Role of intonation patterns in conveying emotion in speech. *ICPhS 1999*, 2001-2004. Retrieved from https://pure.tue.nl/ws/portalfiles/portal/1354072/696624203788951.pdf

Nanjo, H., & Kawahara, T. (2004). Language model and speaking rate adaptation for spontaneous presentation speech recognition. *IEEE Transactions on speech and audio processing*, *12*(4), 391-400. DOI: 10.1109/TSA.2004.828641

Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003, April). Classification of stress in speech using linear and nonlinear features. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on* (Vol. 2, pp. II-9). IEEE. DOI: 10.1109/ICASSP.2003.1202281

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' voices and refusal rates in telephone surveys. *Public Opinion Quarterly*, *50*(1), 97-111. https://doi.org/10.1086/268962

Packwood, W. T. (1974). Loudness as a variable in persuasion. *Journal of Counseling Psychology*, *21*(1), 1. http://dx.doi.org/10.1037/h0036065

Pavlenko, A. (2007). *Emotions and multilingualism*. Cambridge University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=7rt5t2Gzm1wC&oi=fnd&pg=PP1&dq

=Emotions+and+multilingualism.&ots=ALZ2bcnoNK&sig=OnbQhizloMwSwirvq_Hgf4jtnxw#v=onepage&q=Emotions%20and%20multilingualism.&f=false

Pear, T. H. (1931). Voice and personality. Retrieved from https://pure.mpg.de/rest/items/item_2368438/component/file_2591624/content

Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across-language perspective on speech information rate. *Language*, *87*(3), 539-558. doi:10.1353/lan.2011.0057

Protopapas, A., & Lieberman, P. (1997). Fundamental frequency of phonation and perceived emotional stress. *The Journal of the Acoustical Society of America*, *101*(4), 2267-2277. https://doi.org/10.1121/1.418247

Rothkrantz, L. J., Wiggers, P., Van Wees, J. W. A., & van Vark, R. J. (2004, September). Voice stress analysis. In *International conference on text, speech and dialogue* (pp. 449-456). Springer, Berlin, Heidelberg. Retrieved from https://link.springer.com/chapter/10.1007/978-3-540-30120-2_57

Ruiz, R., Absil, E., Gramatica, B., Harmegnies, B., Legros, C., & Poch, D. (1995). Spectrum-related variabilities in stressed speech under laboratory and real conditions. In *Speech under Stress*. Retrieved from https://www.isca-speech.org/archive_open/archive_papers/sus_95/sus5_049.pdf

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408-2412. https://doi.org/10.1073/pnas.0908239106

Scherer, K. R. (1972). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences.

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception 1. *Journal of personality*, *40*(2), 191-210.

Scherer, K. R., Grandjean, D., Johnstone, T., Klasmeyer, G., & Bänziger, T. (2002). Acoustic correlates of task load and stress. In *Seventh International Conference on Spoken Language Processing*. Retrieved from https://www.isca-speech.org/archive/archive_papers/icslp_2002/i02_2017.pdf

Selye, H. (1987). *Stress*. Sterling Educational Film.

Sigmund, M. (2013). Statistical analysis of fundamental frequency based features in speech under stress. *Information Technology and Control*, *42*(3), 286-291. doi: http://dx.doi.org/10.5755/j01.itc.42.3.3895

Smith, B. L., Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Effects of speech rate on personality perception. *Language and speech*, *18*(2), 145-152. https://doi.org/10.1177/002383097501800203

Solé, M. J. (2007). Controlled and mechanical properties in speech. *Experimental approaches to phonology*, 302-321. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=o1MVDAAAQBAJ&oi=fnd&pg=PA302&dq=Controlled+and+mechanical+properties+in+speech.&ots=ShSg-nPgpR&sig=-jOiIzBNkDRArPoN88OMskGGUpA#v=onepage&q=Controlled%20and%20mechanical%20properties%20in%20speech.&f=false

Sorg, B. A., & Whitney, P. (1992). The effect of trait anxiety and situational stress on working memory capacity. *Journal of Research in Personality*, *26*(3), 235-241. https://doi.org/10.1016/0092-6566(92)90041-2

Steeneken, H. J., & Hansen, J. H. (1999, March). Speech under stress conditions: overview of the effect on speech production and on system performance. In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*(Vol. 4, pp. 2079-2082). IEEE. DOI:10.1109/ICASSP.1999.758342

Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. *Journal of personality and social psychology*, *35*(5), 345. http://dx.doi.org/10.1037/0022-3514.35.5.345

Streeter, L. A., Macdonald, N. H., Apple, W., Krauss, R. M., & Galotti, K. M. (1983). Acoustic and perceptual indicators of emotional stress. *The Journal of the Acoustical Society of America*, *73*(4), 1354-1360. https://doi.org/10.1121/1.389239

Šulc, J., & Remek, V. (1986). Possibilities and limits of using speech signals in aviation and space psychophysiology. *Acta neurobiologiae experimentalis*.

Tolkmitt, F. J., & Scherer, K. R. (1986). Effect of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 302. http://dx.doi.org/10.1037/0096-1523.12.3.302

Wu, D., Parsons, T. D., & Narayanan, S. S. (2010). Acoustic feature analysis in speech emotion primitives estimation. In *Eleventh Annual Conference of the International Speech Communication Association*. Retrieved from https://www.isca-speech.org/archive/archive_papers/interspeech_2010/i10_0785.pdf

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology*, *9*(2), 79-94. Retrieved from https://www.researchgate.net/profile/Sanwar_Sunny2/post/What_are_factor_analysis_assumptions_in_using_dichotomous_variables_in_sample/attachment/59d6586579197b80779ae576/AS:538355287166976@1505365310971/download/A+Beginner's+Guide+to+Factor+Analysis+_+Focusing+on+Exploratory+Factor+Analysis.pdf