PERCEIVED TRUST OF UNITED STATES POLITICIANS THROUGH PROSODY

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment
of the requirements for the degree

Master of Science

by

Michael Clemens

August 2020

APPROVAL SHEET

This Thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science

Author Signature: M Clem

This Thesis has been read and approved by the examining committee:

Advisor: Laura Barnes

Committee Member: Medhi Boukhechba

Committee Member: Heather Frasch

Committee Member: Luke Dahl

Committee Member: Jane Hilton

Committee Member: _____

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

August 2020

0.1 Acknowledgements

First, I need to thank Rogelio Cardona-Rivera and the members of the Laboratory for Quantitative Experience Design, particularly Alex and Keith, for their endless support throughout my graduate studies. You all have been instrumental in getting me where I am today. Thank you for the wild information exchanges and for being my academic family.

To my advisor, Laura Barnes, committee member Medhi Boukhechba, and fellow mem-bers of the Sensing Systems for Health Lab - Thank you for welcoming me with open arms. I could not have succeeded in this journey without the encouragement that I received from every individual in the lab.

Thank you to my committee members Luke Dahl, Heather Frasch, and Jane Hilton for always being there to lend an ear and help in whatever way possible. You all provided a great deal of assistance along the way and for that, I am grateful.

To my roommates through the years: Chris, Jonathon, and Joe. Thank you for not only supporting me academically but also emotionally through this journey. From listening to me complain about mundane issues to reading over my paper submissions for the hundredth time, I cannot thank you enough. Thank you for helping create a supportive living environment.

Many thanks to my friends who regularly called or visited to check on me: John, Kyli, Courtney, Helen, Laurin, Mark, Andrew, Bernie, Rick, Jeff, Duncan, Mike, and Tom. From hour-long phone calls to board game nights, you all kept me sane through this journey. There really are no words to describe how much you all mean to me.

I would also like to thank all of the musical artists who got me through this degree: EDEN, Madeon, Skogsra, Kaskade, Oh Wonder, The Strokes, blackbear, Daft Punk, Bad Suns, Whethan, Lemaitre, and many others. Whether or not they know it, these artists are why I got through the toughest portions of this undertaking.

Finally, my biggest thanks goes to Russ. I would not be here today without your endless support. Thank you for making the world a better place.

ABSTRACT

Trust is an essential aspect of interpersonal communication, whereby individuals make instinctive trustworthiness judgments to determine how best to interact with other involved parties. The voice is unique to every individual due to the physiology of both the vocal tract and resonance throughout the rest of the body. When listening to a stranger's voice, humans have developed an innate ability to infer the speaker's trustworthiness. This thesis investigates how individuals determine this perceived trust based on the rhetorical prosodic elements of the speaker. Previous trust literature has focused on three primary contexts for establishing trust: general, mating, and economic. Political trust literature has often used trustworthiness as a means of predicting voter outcomes. This thesis serves to bridge the gap between these areas of knowledge by understanding the level of high-stress situational trustworthiness given to United States politicians based on their prosody. This thesis includes three experiments: Pilot Study (COVID-19 & U.S. Governors), Experiment I (COVID-19 & U.S. Governors), and Experiment II (Protests on Police Brutality & U.S. Mayors). The pilot study (N=141) confirms correlations between prosodic elements of U.S. Governor's voices and their perceived trustworthiness by listeners in both general and high-stress contexts. Experiment I attempts to replicate this correlation with two distinct participant pools (MTurk N=93 & Snowball N=91) using sound bites from eight U.S. Governor's COVID-19 reopening plans. Experiment II also uses two participant pools (MTurk N=93 & Snowball N=88) and expands on the findings from both the pilot study and Experiment I by having participants answer trust-perception questions about eight U.S. Mayors' speeches regarding protests occurring in their respective city. The mayors were selected based on the eight cities with the highest number of police killings from January 2013 through December 2019.

The results from these three experiments strengthen the argument that rhetorical prosodic elements of speech play a key role in determining the speaker's perceived trustworthiness. Linguistic elements, including articulation rate, average syllable duration, and phonation

time, were analyzed; however, results from these studies show that pitch inflections (standard deviation) best explain trust perception. Vocal profiles of prominent historical figures (MLK, JFK, Bill Clinton, and Richard Nixon) were also analyzed and shown to correlate directly with these studies' findings. This thesis demonstrates that rhetorical prosodic elements have a significant influence on trust perception during high-stress situations. These findings can be used by vocal coaches and speech-language pathologists to help clients who are in roles of information dissemination during high-stress situations portray themselves as trustworthy figures.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Trust

Trust is an essential part of initiating and maintaining interpersonal relationships in various contexts, including mating and trading (West et al., 2011). Formal definitions have included "An expectancy held by an individual or group that the word, promise, verbal or written statement of another individual or group can be relied upon" (Rotter, 1967), "The extent to which one is willing to ascribe good intentions to and have confidence in the words and actions of other people" (Cook and Wall, 1980), "An expectation of goodwill and benign intent." (Yamagishi and Yamagishi, 1994), and "The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other party will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al., 1995). From these definitions, trust is a combination of uncertainty, fragility, hope, and expectation. The trustor creates a mental model of how trustworthy the trustee (the individual who is trusted) is and makes a decision based on this model (Gambetta et al., 2000). This is what we define as "perceived trustworthiness," the amount of trust the trustor places in the trustee. Although this model may be wholly inaccurate, this is the risk that the trustor takes to engage in meaningful dialogue.

In the realm of trust literature, researchers make a distinction between the types of trust people exhibit: cognitive and affective. Cognitive trust refers to the beliefs based on the opinions or knowledge about objects (Rosenberg et al., 1960). The level of trust attributed to an object is quickly constructed in this category and is composed of rational content, such as beliefs and intentions (McKnight et al., 2002; Falcone and Castelfranchi, 2001). Cognitive trust ascribes a certain amount of trust to the trustee based on compe-

tence, responsibility, and dependence (Butler Jr, 1991; McAllister, 1995; Lewis and Weigert, 1985). Rather than based on logic and reasoning, affective trust is related to the influential emotional connection provided by the amount of care and concern the trustee shows towards other people (Johnson-George and Swap, 1982; McAllister, 1995). Castelfranchi and Falcone (2010) present this type of trust as well as intuitive forms of trust and refer to trust as a feeling that can be greatly influenced by the emotional appraisal of events.

To help the reader better understand these two concepts, a personality trait is described in detail for each type of trust: reputation for cognitive trust and cooperativeness for affective trust. Misztal (1996) explains that reputation "helps us to manage the complexity of social life by singling out trustworthy people - in whose interest it is to meet promises." Abdul-Rahman and Hailes (2000) define reputation as an expectation held by others based on information about or observations of a person's past behavior. Wang and Vassileva (2003) note that the trustor places beliefs about the skills and honesty of the trustee based upon the recommendations given by others. Since cognitive trust relies heavily upon logic and a reasoned mindset, knowing the experiences, skills, and overall reputation of a trustee can help the trustor ascribe an appropriate amount of trust to the trustee.

Cooperativeness, on the other hand, can be thought of as contributing to the overall affective trust model. Gambetta et al. (2000) state that a person is deemed trustworthy when "the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him" and Burt and Knez (1995) go even further saying that "Trust is anticipated cooperation." Although trust and cooperativeness are often studied together in psychology and voice literature (Krumhuber et al., 2007), they are not the same and can happen without one another, as in the case of children's unconditional trust. Cooperation can occur without trust as well, when the situation presented does not put any of the involved parties at risk (Mayer et al., 1995). This level of cooperation is demonstrated through various peer pressure situations when the risks involved are low. Despite trust sometimes occurring without cooperation and vice versa, it is essential to note that the two are deeply interwoven and contribute to the other.

1.2 Prosody

Prosody is a term used to describe a time series of speech-related information that is not predictable from a lexical window (i.e. word-sized or sentence-sized) applied to the phoneme sequence (Shih and Kochanski, 2002). Broadly, prosody refers to any type of channel for communication, carrying some amount of information that cannot be fully captured by the lexical channel. With this definition, prosody can also refer to hand gestures, facial expressions, and other types of body language since they communicate some amount of information that is not fully transparent in the lexical channel. However, in this thesis, prosody will be used only to describe prosodic elements in the audio domain, not the visual.

Some information that could be carried in this time series include the following (adapted from (Cruttenden et al., 1986; Couper-Kuhlen, 1986; Ladd, 2008; Gussenhoven et al., 2004; House, 2007)):

- Intonational phrasing: the division the spoken text into chunks
- Accentual highlighting: the distribution of prominence across the text
- Pitch contours or targets: the specification of particular pitch patterns, such as rises or falls
- Pitch span or range: wide or narrow, local or global
- Tempo: fast or slow
- Pauses: both filled and unfilled
- Rhythm: the regular occurrence of speech events in time
- Voice quality : modal, creaky, breathy, and so on

Although the effects of prosodic elements in speech vary widely and are context-dependent, three of them have been widely agreed upon in the literature (adapted from House (2007)):

- Prosodic elements on comprehension can range from 'natural' effects (i.e. an angry, friendly, or grumpy tone) to linguistic (i.e. lexical stress or lexical tone).

- Prosodic elements combine and interact with elements from other sources forming different meanings based on the context. The same prosodic inputs may have vastly different effects based on varying contexts.

- Prosody is often referred to as the 'packaging' of a sentence rather than the 'content' of the message itself. Prosody creates impressions, conveys information about emotions or attitudes, or alters the meaning of language.

It is outside the scope of this thesis to provide more context in prosodic elements. Should the reader want to know more about this subject area, they are encouraged to refer to Cruttenden et al. (1986); Couper-Kuhlen (1986); Ladd (2008); Gussenhoven et al. (2004).

1.3 Summary

This chapter provided a brief introduction to the concepts of trust and prosody. We explored the different elements of trust, cognitive and affective, and discussed what types of prosodic elements play key roles in the comprehension of a sentence. In the next section we will review the literature on trust and political prosody.

CHAPTER 2

LITERATURE REVIEW

2.1 Voice and Trust

In this literature, the halo effect becomes important to note since a good impression created in one area (such as attractiveness) could influence a person's judgement in another (perceived trust) (Nisbett and Wilson, 1977). For example, Zuckerman and Driver (1989) found that attractiveness elicited higher traits of dominance, likability, and even created a "what sounds beautiful is good" stereotype based on these findings. Attractive speakers were also reported as having better job performance, performing better in job interviews, and being more persuasive, competent, and sociable (Burgoon et al., 1990; DeGroot et al., 2011). The issue with these studies is that subjectivity clouds the definition of an attractive voice. Are there qualities that differentiate an attractive voice from an unattractive voice, and will attractive voices always be rated higher in trust than less attractive voices? From a few preliminary studies in the 90s, it was found that an attractive voice was "intermediate in its loudness and more resonant", had faster speech rate, fewer pauses, lower pitch, and greater pitch variability (DeGroot and Motowidlo, 1999; Zuckerman and Miyake, 1993). A recent study on the acoustic features of vocal attractiveness showed the males were rated as more attractive when they had lower first formant frequencies for the "i" and "u" vowels and when the duration was shorter (Babel et al., 2014). The females, on the other hand, were most attractive with a breathy voice and showed "u-fronting" (a trait in many young California women), suggesting group dynamics and stereotypes play into attractive voices. This finding also suggests the role that accents play in determining attractiveness and thus trust. More on that will be discussed in the next section.

Sexual orientation plays a large role in determining attractiveness since the main pur-

pose of attraction is to find a mate (Grammer et al., 2003; Hughes et al., 2004). Females rate low-voiced males as more attractive, while males rate high-voiced females as more attractive (O'Connor and Barclay, 2017; O'Connor et al., 2014). The "breathiness" found in women's voices has been linked to desirability and is rated as more attractive (Henton and Bladon, 1985; Babel et al., 2014). For males, the clear and resonant tone presents dominance that has been linked with attractiveness. Weaker voices are often not attractive since they can be a sign of poor physical health (Blood et al., 1979).

2.2 Biology of Trust

Although trust develops over time as more information about the other party presents itself, several studies show a physiological and neurological response to trust. Through the oxytocin receptor gene, Riedl and Javor (2012) suggested that trusting behaviors are genetically predetermined to some extent (Reuter et al., 2009). The levels of oxytocin were studied during a simulated trust game and found to be higher when participants trusted their partners and when their partners trusted them (Zak et al., 2005). Another group performed the same study, although with nasally administered oxytocin, and found that the group with the oxytocin treatment demonstrated increased trusting behaviors (Kosfeld, 2007).

In terms of neurophysiology, the part of the brain that determines trust and cooperativeness is suggested to be the amygdala. Several brain lesion studies found that patients with damage to the amygdala tended to trust untrustworthy partners in trust games and rate faces as more trustworthy overall than neurotypical individuals (Koscik and Tranel, 2011). A meta-analysis in this field by Bzdok et al. (2011) found that the amygdala was mentioned in numerous studies and suggest that the fight or flight response as well as social behaviors, such as trustworthiness, might be determined by the amygdala. Riedl and Javor (2012) reviewed the literature on neurophysiology and trust and mapped the anatomy to the different aspects of a social context concerning trust:

- reward processing: striatum and thalamus
- risk: amygdala, insular cortex, hippocampus, and parahippocampus gyrus
- memory: necessary for weighing trust at each interaction based on previous experi-

ence — the amygdala, hippocampus, and parahippocampus gyrus

- processing of cognitive conflict: cingulate cortex
- mentalizing: frontal cortex

When meeting a new person for the first time, more often than not, many people will assume they can be trusted since it is both easier and more statistically accurate than to distrust them initially (Jones and George, 1998; Barclay, 2008). Some researchers have found that an individual's mood or attitudes will affect their level of trust towards another individual (Jones and George, 1998). Other research has shown that individual differences play an important part in determining the level of trust given to other people, shown in the way that some people start with the opposite trust level from the norm, meaning the trustor assumes that no-one can be trusted. This mentality is often associated with Machiavelli, who argued that "deceptive behavior is an effective political tool" (Repacholi et al., 2003). These "Machiavellians" portray different neurophysiological responses when associated with trust games than the norm. Machiavellians demonstrate less activation in regions linked to social empathy (medial and ventromedial prefrontal cortices, the inferior parietal cortex, and the superior temporal sulcus). Instead, they activate areas associated with reward-related decision making and inference making (inferior and middle frontal gyrus, the anterior insula, the thalamus, and the anterior cingulate cortex) (Bereczkei et al., 2013).

2.3 Evolutionary Trust

From the studies relating trust with neurophysiology, it is safe to assume that trust also has evolutionary ties in our society. For society to function, trusting decisions are made continuously. We trust our country to act in a way that represents our societal beliefs, protects citizens from harm, provides access to healthcare and education, and defends us from malicious parties. For citizens, we trust one another to follow the laws our country has put in place. These levels of trust are intertwined and necessary for a functioning society (Rothstein and Uslaner, 2005).

Trust is not unique to the human species, but everyone in the animal kingdom benefits if there is trust between individuals (Packer, 1977). "Free riders," people who do not

cooperate while everyone else does, have emerged from this evolution in cooperative contexts, but these individuals benefit only if there are few of them in society (Axelrod and Hamilton, 1981; Doebeli et al., 2004; Trivers, 1971). If no one in the society cooperates and participates then everyone will suffer together. Thus, being able to identify these uncooperative individuals is an advantage for everyone in the group (Doebeli et al., 2004).

Research studies have shown that individuals can distinguish trustworthy from untrustworthy individuals rapidly. In one study, participants were able to develop first impressions of trustworthiness, competence, likeability, attractiveness, and aggressiveness within 100ms (Willis and Todorov, 2006). The main finding, however, that's relevant to trust is that out of all the traits taken into account, trustworthiness was the impression with the highest correlation between judgments with and without the time constraint. The conclusion from the authors was that immediate trust, in evolutionary terms, might have been necessary for survival. Another study, in much the same vein, found that positive traits (extraversion and agreeableness) increased accuracy with exposure time, whereas other traits (negative affect, neuroticism, openness, and intelligence) did not. These findings suggest that the positive traits were more linked to social interactions, which take time to construct a more holistic narrative about the individual, whereas negative traits would be linked to threat and competence for which a quick decision might be warranted. Thus determining untrustworthiness may be extremely important in an evolutionary sense, whereas learning just how trustworthy someone is, is not as pertinent.

To help justify this argument, it has been shown that people remember untrustworthy participants in a dilemma game more than cooperative players (Yamagishi et al., 2003). Another study found that recall for untrustworthy faces was higher than recall for trustworthy faces even though nothing about trustworthiness was mentioned during the recall (Rule et al., 2012). This research aligns with the notion that most people are deemed trustworthy, and only the untrustworthy ones need to be remembered to avoid them in the future (Jones and George, 1998). In much the same light as the Machevillians, a research study found that in artificial societies where deceivers outnumbered cooperators, people remembered cooperators better (Barclay, 2008). Since most people tend to cooperate in society, humans have developed an innate sense of determining and remembering untrustworthy behavior.

2.4 Trust in Personality Traits

Trust is expressed in the agreeableness category for the ”Big 5” with the other four being openness, conscientiousness, extraversion, and neuroticism (McCrae, 2009; Cervone and Pervin, 2015). Broadly, agreeableness can apply to the social perceptual Theory of Mind (ToM) (Nettle and Liddle, 2008). This model represents the way individuals infer someone's mental state based on sensory information such as eye gaze, facial, and vocal expression, rather than social-cognitive ToM, which uses reasoning about someone's mental state to predict behavior (Amodio and Frith, 2006; Tager-Flusberg and Sullivan, 2000). Since trust develops as more evidence is gathered, we can assume that our initial attitudes towards others are based on our social perceptual ToM, meaning that it's constructed through things such as physical traits, behavior, and voice characteristics. Once more information is gathered about the individual and a deeper understanding is beginning to emerge, social-cognitive ToM is used to alter the initial trust decision.

Sometimes, individuals determine trust based on the halo effect where an individual who is good and demonstrates some positive traits will be assumed to demonstrate other positive traits and vice versa (Nisbett and Wilson, 1977). A confounding variable in most trust studies is physical attractiveness since this trait is deeply connected to trustworthiness, making it difficult to compare one without the other (Todorov et al., 2008). As previously mentioned with the amygdala, Bzdok et al. (2011) found that this region of the brain helps determine trustworthiness as well as other long-term social information (”I might want this person to be a future partner”).

2.5 Accents

The definition of an accent is any systematic difference in pronouncing the sounds of a language that people belonging to a certain group share (Lippi et al., 1997). Dialects, on the other hand, are systematic differences in terms of morphology, syntax, lexicon, and pronunciation of a language (Trudgill et al., 2000). From this definition, every speaker has an accent. The media takes advantage of the fact that accents can be linked to personality stereotypes and choose to include native and non-native accents to present a villain or for some type of humor (Dobrow and Gidney, 1998; Preston, 1999). One study showed that native accents might be perceived as more trustworthy than non-native accents (Lev-Ari

and Keysar, 2010).

Accents can determine an individual's nationality, but other vocal characteristics can help a listener determine the person's physical appearance, age, gender, personality, emotional state, and even sexual orientation (Mack and Munson, 2012). A tremendous amount of information is conveyed through voice since it is our primary mode of communication, and research has shown that people are able to decode this ”hidden” information from speech in a short time (McAleer et al., 2014). Duration increases accuracy in personality judgments, which shows that vocal characteristics carry information about personality traits, and we are able to decode this information quickly and accurately. Brown et al. (1975) saw that speakers with a slow speech rate were rated as less competent than faster speakers. Apple et al. (1979) found in a similar study that speakers with slow speech rate were rated as ”less truthful” and ”more passive.” Other studies found that a fast speaking rate is a feature of charismatic (Jiang and Pell, 2017), confident (Hirschberg and Rosenberg, 2005), and persuasive speakers (Chaiken, 1979). The fundamental frequency (F0) of a voice, pitch, is by far the most researched subject within the vocal literature. F0 is the rate of vibration of the vocal folds in the larynx and is partially determined by the size of an individual's vocal folds: large vocal folds produce a lower pitch and vice versa. The literature behind this trait and trust is varied and one can find conflicting results. Tigue et al. (2012) found in a mock election scenario that participants voted for male candidates with lower-pitched voices. Banai et al. (2017) saw over a number of real election results that low pitch was an accurate predictor of the result. Apple et al. (1979) demonstrated that speakers with high F0 were rated as ”less truthful.” On an opposing note, female candidates with high pitch voices were more successful than those with a low pitch when regarding political trustworthiness (Klofstad, 2016). Bonein and Serra (2009) found in an economic game context that participants consistently selected the higher-pitched individual as the person they would trust more with their investment. In studies of deception the literature is contradictory, saying that individuals raise their pitch when lying, a lower pitch is associated with deceptive messages, and one study failing to find any acoustic differences in deceptive and truthful messages (Kirchhübel and Howard, 2013; Villar et al., 2013). Although there is a physiological difference in the vocal structure of females and males, there are no consistent differences in trustworthiness measurements

for either (Bonein and Serra, 2009; Slonim and Guillen, 2010). In one study, although females were rated more trustworthy than males, there were no significant differences found in the trusting behaviors of males and females (Chaudhuri et al., 2013). In this study, gender differences diminished when groups played the game rather than individuals. The last main trait that has been expressed and studied in the literature is voice quality (shimmer, jitter, breathy, etc.) (Klatt and Klatt, 1990). Laver (1968) found that harsh voices correlated more towards aggression and dominant characteristics, whereas breathy voices were characterized with submissive personalities. Blood et al. (1979) found hypernasal and breathy voices were rated more negatively than the control. Breathiness was found to be attractive in both males and females because it implied a reduction in aggression (Xu et al., 2013; Gussenhoven, 2016).

2.6 Political Prosody

In politics, the rhetorical style used is often done in a directly targeted manner. The persuader, the politician, uses target-oriented rhetoric specifically for their audience. Touati (1991) explored how some of these rhetorical features translated into prosodic elements in the context of a pre-electoral television debate. In a follow-up study, Touati (1993) found that pitch variation is used when establishing a paradigmatic contrast between pre- and post-electoral speech or a change between spontaneous speech and rehearsed speech (i.e. when directly quoting someone). Register change was also used when quoting the opposition. Touati's work provided the foundation for future research of political prosody to investigate both register and pitch variation while conducting rhetorical analysis. Recent political literature looks to these prosodic elements as a way of predicting voter outcomes.

Voting behavior involves a complex interaction of factors, but some researchers suggest that mate-related choices and attractiveness can help predict constituent voting decisions. Navarrete et al. (2010) showed that women's conception risk across the menstrual cycle positively predicted their intention to vote for Barack Obama in the 2008 US presidential election. This effect was shown to be strongest among women who perceived his voice as more indicative of a white candidate than black. Little et al. (2007) showed that constituents preferred to vote for candidates with relatively more masculine and dominant faces but not more seemingly "attractive faces." In this study, Little also showed that voters

preferred the dominant faces during wartime but the more ”attractive faces” in times of peace.

Rendall et al. (2007) published a study stating that individuals could accurately predict the upper body strength of a male based solely on men's voices and vocal cues. Unlike body size, which has a reputation of being predicted incorrectly, the ability to accurately estimate the strength of an individual allows the listener to make other predictions based on that quality. Some studies in this realm show that voter's preference is due in part from the vocal attractiveness of the speaker. This attraction leads people to believe the speakers were also more trustworthy as well as other positive attributes (Tigue et al., 2012; Surawski and Ossoff, 2006). This was found to be true in both explicit and implicit measurements of behavior (Chaiken, 1979).

Recently, more studies have been published that focus once again on using voice to predict voter outcomes. For example, a number of studies using experimental and real-world electoral data found that candidates with lower-pitched voices win more votes (Tigue et al., 2012; Anderson and Klofstad, 2012; Klofstad, 2013; Klofstad et al., 2012). Researchers have argued that since deeper voices have been shown to be associated with strength, dominance, muscularity, and masculinity, (Evans et al., 2008; Feinberg et al., 2005; Puts et al., 2006; Sell et al., 2010) this preference provides evidence for the preference of deeper voiced leaders as they are looked upon to protect and prevail in times of war and conflict (Tigue et al., 2012). A follow-up study by Laustsen et al. (2015) found that conservative Republicans preferred lower-pitched voices than do liberal Democrats.

2.7 Physiology of the Voice

The voice consists of three individual pieces that work in tandem: voiced sound, resonance, and articulation.[1] Voiced sound refers to the sound produced by the vibration of the vocal folds. Resonance amplifies the voiced sound and is modified by the vocal tract resonators (the throat, mouth cavity, and nasal passages). At this stage of the sound creation, the sound produced is the individual sound emitted from a person. The last piece, articulation, consists of the tongue, soft palate, and lips and modifies the voiced sound to

[1] https://voicefoundation.org/health-science/voice-disorders/anatomy-physiology-of-voice-production/

produce recognizable words. The way in which the vocal folds vibrate gives the speaker their individual texture and formants. Formants of a voice refer to the characteristic resonance region that is dictated by the individual's size and shape of their vocal tract, nasal cavity, and oral cavity.[2] This is why when someone gets sick, their voice changes.

For the sake of brevity many details of vocal production have been omitted, but hopefully this provides the reader a brief introduction into the physiology of voice production and will serve as a reference as we start exploring the various experiments presented in this thesis.

2.8 Summary

This chapter provided relevant literature on trust, political prosody, and the voice. We learned about what elements of the human body make up vocal production, what biological and psychological traits make us trust (or distrust) others, and what elements of prosody have been used to predict constituent voting behavior. To the best of the author's knowledge, no study has looked in-depth at the type of relationship that prosodic elements have towards trust. Most literature in this realm has assumed a linear relationship between these components, while the results shown in this thesis demonstrate that more complex interactions may exist. In the next chapter, we will explore three experiments used to help determine what prosodic elements constituents use to decipher the amount of trust they should give to United States politicians during times of crisis and the various complexities of these relationships.

[2]https://www.sfu.ca/sonic-studio-webdav/handbook/Formant.html)

CHAPTER 3

EXPERIMENTS

3.1 Pilot Study: COVID-19 & U.S. Governors

3.1.1 Introduction

The COVID-19 (previously known as "2019 novel coronavirus") pandemic has consumed world news in unprecedented ways and shattered healthcare, economic, and social norms in the United States. Presented below is a brief timeline to illustrate just how fast the virus spread and the delay in government reaction. From the World Health Organization's (WHO) information regarding the novel coronavirus pandemic[1]:

- 2019/12/31 - WHO is notified about a cluster of pneumonia cases of unknown etiology in Wuhan, China
- 2020/01/07 - China reports the virus to be a new type of coronavirus
- 2020/01/11 - China shares the genetic sequence
- 2020/01/30 - WHO declares a Public Health Emergency of International Concern
- 2020/01/31 - 9,826 cases of COVID-19 in 20 countries and 213 deaths (9,720 of those cases in China)
- 2020/02/29 - 85,403 cases of COVID-19 in 54 countries. 2,924 deaths in 7 countries. 93% of cases and 97% of deaths in China.
- 2020/03/11 - WHO classifies COVID-19 as a pandemic

[1]https://www.who.int/images/default-source/departments/epi-win/infodemic-management/infodemic-management-covid19.jpg?sfvrsn=51e4edb8$_4$

- 2020/04/16 - Mr. Trump's coronavirus task force released broad guidelines for states to reopen in three phases based on the number of cases within each state and the hospital capacity.[2]

- 20/05/20 - The Centers for Disease Control and Prevention (CDC) quietly released their reopening guidelines, initially deemed too rigorous and limited by the Trump administration.

On that same day in May, CNN released an interview with several CDC employees explaining the contentions between the White House and the CDC regarding reopening plans and whose opinion should dictate policies. These arguments started back in February and have continued since then regarding policy declarations.[3]

The arguments surrounding data and government are not limited to the federal government. The Georgia Department of Health was heavily criticized for manipulating bar charts on their website's COVID tracker to make it appear as if cases were going down, rather than reporting accurate data showing that the case numbers were escalating.[4] When a top data scientist in Florida refused to manipulate information on case numbers, to appear that cases were taking a downward trend, she was fired. After leaving the team, she published a statement regarding her concerns that Florida's COVID tracker data is being tampered with and displaying misleading information to the public.[5]

In 1984, Britain passed the "Police and Criminal Evidence Act," which altered how police interactions and interviews were administered.[6] Rather than police merely listening to the story, writing up a police report, and having both parties sign the report, the entire exchange was now to be recorded via audio recorders. This act was put in place to stop potential verbaling (putting words into people's mouths) from occurring during police

---

[2]https://www.whitehouse.gov/wp-content/uploads/2020/04/Guidelines-for-Opening-Up-America-Again.pdf

[3]https://lite.cnn.com/en/article/h$_{9}$c8403ed93f40b6a45c60ae47e1be0d2

[4]https://www.cnn.com/2020/05/20/us/florida-georgia-covid-19-test-data/index.html

[5]https://www.npr.org/sections/coronavirus-live-updates/2020/05/19/859119865/florida-ousts-top-covid-19-data-scientist?utm$_{m}edium = RSSutm_{c}ampaign = nprblogscoronavirusliveupdates$

[6]http://www.legislation.gov.uk/ukpga/1984/60/contents

interviews to help mitigate wrongful convictions taking place. However, this act turned out to be more useful than the government had originally intended because it carved the way for the art of ”forensic listening.” While the government was interested in what people said, it was rather the sonic textures in the environment that provided more clues to the police’s investigation than any of the semantic content that was being exchanged. Police began calling on linguists and phonologists to decipher if the background noise in a recording was a gunshot or someone slamming a door. This legislation initiated the movement that governs our speech today.

Lawrence Abu Hamden calls 1984 the avant-garde of listening because it sparked a new birth in the era of ”forensic phonetics.”[7] Before the ”Police and Criminal Evidence Act” was voted in, it was expected that people would undergo a transformation when they stepped onto a witness stand, and speech would transcend from ordinary conversation to liable testimony. Now with the recordings, however, all speech and sound are liable wherever we are. Our speech is legally accountable in all places, no longer masked behind a paper’s accounts of our words. It is not only what we say that matters, but how we say it and where that also determines how people interpret our speech.

With the amount of misinformation being spread about COVID-19 online and within our own friend groups, who can we trust? All of our speech (be it audio or through text) is liable to whatever government entities we are subject to. Our federal and state governments are fighting with health experts and data scientists regarding reopening policies and case counts. Which numbers are correct? Should these individuals be liable for deliberately falsifying data? Why do we pick a particular news station to listen to for the current updates? Do we trust the source because of the name, because we grew up listening to that style of rhetoric, or is it something completely different? This experiment was designed to determine what intrinsic audio or sonic cues United States residents listen for to determine the speaker’s perceived trustworthiness.

---

[7] https://new.bidoun.org/articles/lawrence-abu-hamdan

3.1.2 Method

3.1.2.1 Participants

There were a total of 161 participants in this study gathered by Snowball sampling. Posts were made on Twitter, Facebook, and LinkedIn to gather a more diverse and representative sample. The only requirement for participating in the survey was that the participant answered all the required questions and reached the final screen before terminating the website. 20 participants did not reach the end of the survey and were removed from the analysis. Of the 141, 61 were male, 49 were female, and four were non-binary. 20 had a high school diploma or GED, 50 had a Bachelor's, and 30 had a Masters's degree or higher. Eight were Republicans, 53 were Democrats, 29 were independents, and 10 did not define themselves as affiliated with any party. 53 participants were between the age of 18-29, 13 were between 30-39, 11 were between 40-59, and 24 were 60 or older. Five made less than $20,000 a year, 34 made between 20-49k, 24 made between 50-74k, 17 made between 75-99k, and 16 made equal to or above 100k. Participants resided in 20 different states throughout the United States, although there was a heavy leaning in Indiana, Virginia, and Utah.

3.1.2.2 Stimuli

All 50 United States Governors' voices were used in this study. Recordings were gathered from YouTube, Governors' websites, and news agencies' websites. The video clips were played back through a Komplete Audio 6 (Audio/MIDI Interface developed by Native Instruments)[8] and recorded in Ableton Live 10.[9] Some of the clips were only uploaded to their respective distribution service using one instead of two channels of audio; thus, while recording, each track was converted to mono instead of stereo. Every sound bite was processed using a de-clipper, noise remover, de-esser, plosive remover, and compressor in that order. These virtual plugins were from an audio production suite created by Accusonus.[10] This processing was necessary to ensure that each sound bite

[8] https://www.native-instruments.com/en/products/komplete/audio-interfaces/komplete-audio-6/

[9] https://www.ableton.com/en/

[10] https://accusonus.com

was as close to one another in terms of background noise, loudness, and clarity to reduce the possibility for confounding variables significantly impacting the results of the survey. Each clip was between 30 and 45 seconds long and included semantic content regarding their respective state's COVID numbers and reopening procedures. Each clip was also gathered between April 20th and April 24th, which was after The White House released their guidelines on reopening procedures for the states.[11] These dates are important to note since each state Governor had ample time to discuss their reopening procedures and come up with a viable plan for the future. The sound bites were gathered in the first press conference for each state after the guidelines were announced, and a plan for each state began to materialize.

3.1.2.3 Procedure

Participants were given a link to a survey and completed it at their leisure with whatever audio equipment they use daily. The first question posed was about state residency. The answer to this question was used to curb potential bias towards their own Governor. That Governor was then removed from the list of potential sound bites. Participants were then instructed to rate six randomly selected sound bites (three female and three male) on three questions:

1.) "How trustworthy does the speaker sound?"
2.) "How likely is it that this person will provide you with factual information about the COVID-19 pandemic?"
3.) "How likely is it that this person will provide you with factual information during a natural disaster?"

A 7-point Likert scale was used to answer each question with (1-very untrustworthy to 7-very trustworthy) for question 1 and (1-very unlikely to 7-very likely) for questions 2 and 3. These questions and 7-point Likert scale were adapted from Schild et al. (2020). Demographic questions were asked at the end with a heavy leaning on political thoughts and opinions. The full list of questions is provided in Appendix C.1.

[11] https://www.whitehouse.gov/openingamerica/

3.1.3 Results

All sound bites were analyzed using Praat software (v 6.1.13) (Boersma and Van Heuven, 2001). Before any analysis was performed, the sound bite was assessed for background noise and extraneous pitches that could skew results. Each sound bite was converted from the time domain into the frequency domain with the "To Pitch" command in Praat. This command performs an acoustic periodicity detection from an autocorrelation method. This method is "more accurate, noise-resistant, and robust than methods based on cepstrum or combs, or the original autocorrelation method" (Boersma, 1993). Since some sound bites were taken from less than optimal recording samples, some adjustments to the settings were made to help fine-tune what the algorithm was reacting to. The time step remained the same as per the programmer's recommendations (0s). However, the pitch ceiling was changed from 600Hz to 300Hz and the pitch floor was altered from 75Hz to 50Hz. Once each pitch profile was created, the author went through each profile and either unvoiced some pitches that were not the Governor's voice (i.e. someone screaming from the crowd or the plosive "pop" from the microphone) or changed the pitch to one of equal value that was appropriate. For example, if the Governor had a breathy tone, a whistle tone may have been picked up by the algorithm, indicating a pitch of much higher value than the Governor's vocal tone. Only comparable numbers (9 to 9 or 1 to 1) were used to alter these individual pitches. Three parameters of fundamental frequency were acquired using Praat's "voice report" feature: mean F0 (Hz), range (Hz), and standard deviation (Hz). For other aspects of prosody, the Praat script "Syllable Nuclei" was used (De Jong and Wempe, 2009). The features captured by this script include speech rate, articulation rate (nsyll/phonation time), average syllable duration (speaking rate/nysll), phonation time, and duration.

These voice profiles were analyzed after the study was complete. Rather than running all 50 Governor's voices through the software initially, the author waited for the results to dictate which Governor should be looked at in-depth. The reader should notice here that the list of Governors in Table 3.1 and Table 3.2 do not match. Table 3.1 includes the frequency distributions of the Governors that were chosen to move onto the second round, as discussed by the credentials above. Table 3.2, on the other hand, shows the voice profile analysis of each Governor that was clearly perceived as high or low trust based on the

results from the survey. The only Governor from Table 3.1 that is missing from Table 3.2 is the Governor of Iowa, since she was still voted ”high-trust”, but was second-lowest overall for the female Governors.

		AL - 1	Iowa - 15	MI - 22	SD - 41	AK - 2	ID - 12	VT - 45	WI - 49
General Trust	1	2	1	1	1	0	0	0	0
	2	7	4	1	1	4	1	0	0
	3	7	7	2	0	0	3	1	0
	4	7	4	3	3	1	2	1	0
	5	5	12	10	6	2	3	1	2
	6	8	9	12	15	2	0	4	3
	7	1	3	10	9	0	1	3	3
	Trust Ratio	-0.05	0.30	0.72	0.80	0.00	0.00	0.70	1.00
COVID Trust	1	2	2	1	1	0	0	0	0
	2	9	4	0	0	3	2	0	0
	3	10	3	3	2	2	2	1	0
	4	6	4	3	3	0	3	1	0
	5	6	12	8	5	2	3	2	2
	6	4	11	14	13	2	0	4	3
	7	0	4	10	11	0	0	2	3
	Trust Ratio	-0.30	0.45	0.72	0.74	-0.11	-0.10	0.7	1
Disaster Trust	1	3	2	1	0	0	0	0	0
	2	1	3	0	1	1	2	0	0
	3	7	5	3	3	4	2	1	0
	4	9	2	2	1	0	3	1	1
	5	7	12	9	6	1	3	1	2
	6	4	13	13	18	3	0	3	2
	7	6	3	11	6	0	0	4	3
	Trust Ratio	0.16	0.45	0.74	0.74	-0.11	-0.10	0.70	0.88

Table 3.1. Frequency Distribution for the pilot study. Red columns indicate low perceived trust; blue columns indicate high perceived trust.

	Female Gov			Male Gov							
	AL-1	MI-22	SD-41	AK-2	CA-5	HI-11	ID-12	MT-26	NC-33	VT -45	WI-49
notes	low	high	high	low	high	high	low	low	high	high	high
mean F0 (Hz)	195.69	197.97	194.51	144.71	137.15	138.49	133.19	143.79	135.29	156.38	140.24
range F0 (Hz)	217.50	229.60	112.80	237.10	73.35	98.78	159.40	244.10	260.80	170.40	120.30
std dv F0 (Hz)	32.13	37.61	18.32	31.44	12.75	19.19	30.41	32.92	37.06	22.89	18.37
speech rate (nsyll/dur)	4	3.36	4.62	4.13	3.5	3.46	2.72	3.05	2.41	3.82	3.72
articulation rate (nsyll / phonationtime)	5.12	4.47	5	5.2	3.86	3.86	4.35	4.48	3.85	4.65	4.31
ASD (speakingtime/nsyll)	0.195	0.224	0.2	0.192	0.259	0.259	0.23	0.223	0.26	0.215	0.232
phonation time	28.7	30.41	34.18	31.73	37.53	37.86	23.46	29.92	24.67	36.55	33.66
dur (s)	36.71	40.48	36.98	39.98	41.47	42.23	37.47	43.98	39.46	44.47	38.96

Table 3.2. Praat analysis of governors in pilot study.

3.1.3.1 Trust Measures

Since the data recorded were Likert-type data, instead of Likert-scale data, the mean would not be appropriate to gauge how trustworthy a Governor was perceived. Instead, a frequency distribution was used and is provided in Table 3.1. The Trust Ratio (number of positive responses to negative responses) is also shown to gauge the number of people who thought the speaker was trustworthy or untrustworthy. Depending on the literature, the mean of Likert data can be used as a telling metric. However, in this study, if the mean response is said to be 2.5, does that suggest the person is more untrustworthy than "somewhat untrustworthy" but not enough to be "untrustworthy?" In light of this argument, the trust ratio is used throughout this paper to measure how trustworthy a speaker is perceived.

This study was purely speculative and used to determine if any correlations may have existed between the speaker's prosody and their perceived trustworthiness. As such, no statistical tests were performed (i.e. t-tests or Pearson's r). This study is repeated in Experiment I to validate the claims found through this pilot study and determine the statistical significance of these correlations.

3.1.4 Discussion

The results from this study show a clear correlation between perceived trustworthiness and prosodic elements. Based on the demographic information provided, there was a normal distribution of perceived trustworthiness for age, finances, and education. In context, the youngest and oldest participants voted everyone lower overall, whereas people in their 30s-50s voted Governors as higher trust. People who made between $50-74k per year voted Governors higher trust than people who made more or less than that. People who had a Bachelor's degree rated Governors higher trust than people with more or less formal education. Female Governors were voted higher trust overall than male Governors (but since there were only nine females and 41 males, this statistic is not as telling as it might appear). Heterosexuals voted Governors higher overall as compared to people who identify as non-heterosexual. Females voted for Governors significantly higher in trust overall than men did.

From the results of the study, the Governors from Table 3.2 all went through a barrage of

vocal processing techniques to determine what prosodic features correlated directly with perceived trustworthiness. First, some musical information retrieval techniques were used to determine if the energy in the voice (albeit very minute) could be differentiated from one Governor to the next. Zero Crossing Rate and Spectral Centroid were used, but showed no consistent correlations. Mel-Frequency Cepstrum Coefficients (MFCCs) were calculated for each sound bite and examined for any possible correlations. The only correlation from this endeavor was that the large gaps between changes in energy were from a lack of dialogue. This was investigated further in both Experiments to ensure that phonation time (the amount of time the person spoke during the sound bite) was not directly correlated to how trustworthy the speaker was rated.

Mean F0, range, standard deviation, speech rate, articulation rate, average syllable duration, phonation time, and duration were all studied for possible correlations. To reiterate, since so few voters ranked male Governors, it would be useless to perform statistical tests on these correlations. Even if statistical significance was demonstrated, the actual significance of these correlations would be absent. The three metrics that seemingly had correlations were articulation rate, phonation time, and standard deviation with the last one being the most correlated with the perception of trustworthiness. Experiment I takes these findings and uses them as hypotheses in its testing.

There was found to be little to no difference between the responses in question 2 and 3 (COVID trust and natural disaster trust). Therefore, question 3 was eliminated from Experiment I.

3.2 Experiment I: COVID-19 U.S. Governors

3.2.1 Introduction

This study is a continuation of the pilot study, using the same sound bites, and attempting to draw similar correlations and conclusions. From the 50 Governors that were used in the pilot study, only eight (four female and four male) Governors were used. The selection criteria was a mix between the frequency of each response, the mode, and the trust ratio (the ratio between positive and negative responses for each Governor).

3.2.2 Method

3.2.2.1 Participants

This study's participants were split into two distinct participant groups: Snowball sampling and Amazon Mechanical Turk (AMT). In both groups, participants were required to be 18 years or older and a current United States resident. These two groups were included to ensure that the results were from a more heterogeneous sample than either population independently. The author compensated participants through AMT from his personal account. The compensation for the survey was based on the Qualtric's informed completion time for the survey (5 mins) and the federal minimum wage in the United States ($7.25). Therefore, each participant in the AMT population received $0.60 for completing the survey. There was no compensation provided for the Snowball sample.

The sample size was based on the numbers in Krejcie and Morgan (1970), where N is the population of the United States (328M). With this N, the number of participants necessary for this study is 380. Research by Kordsmeyer et al. (2018); DeBruine and Jones (2018); Hehman et al. (2018) has shown that about 15 voice raters were able to replicate these results for each category they targeted. In the current study, raters were asked two questions regarding each Governor; thus, 30 participants would be necessary to achieve the effects described in the previous studies. In Schild et al. (2020), the researchers used 15 females and males for each trust context, implying that there should be 30 participants for each measurement. To achieve a representative and heterogeneous sample, two population pools were used and a minimum of 100 participants for each group's sample size. With the additional participants in Experiment II, the number of participants was closer to the sample size as stated by Krejcie and Morgan (1970) (N=380).

104 participants in total accepted the Human Intelligence Task (HIT) on AMT. To be considered for the HIT, the turker had to have one HIT approved and a HIT approval rating for all requesters' HITs greater than 95%. These requirements go directly against standard research practice with AMT, but with the COVID pandemic leaving people out of a job or working from home, some individuals are turning to AMT as a source of income.[12] Some researchers have condemned using AMT because there may be data

[12]https://www.wired.com/story/newly-unemployed-labeling-photos-pennies/

quality issues, leaving researchers to vet these turker's responses to guarantee that high data quality is preserved when completing their study. However, with approval rate inflation, non-naivety of "superworkers", and a vast majority of U.S. workers new to the platform each year, it becomes a question as to why researchers are set on using this platform in a homogenous manner, thereby indirectly polluting their "representative" sample (Robinson et al., 2019). This present study deviates from the norms and attempts to gather "naive" workers who are just starting on AMT to help both support their initiative during the COVID-19 pandemic and achieve a more heterogeneous sample. Out of the 104 participants, one entered the wrong survey code at the end, three responded to the attention checks incorrectly, and six were under the 180 second time requirement. This time requirement was put in place to eliminate respondents who might have just clicked through the survey without fully listening to the samples. The minimum time to listen to the samples took 120 seconds, and 60 seconds were added for individuals to read the directions and answer the various questions. Turkers also only had an hour to complete the HIT. After these requirements, 94 participants were left from the AMT population pool.

The Snowball sample consisted of participants from Twitter, Reddit, and LinkedIn. Overall, there were 133 survey responses from participants. 33 participants did not fully complete the survey and were eliminated from the analysis. Four participants answered the attention checks incorrectly, three were not United States residents, and two participants took over 3600 seconds. These requirements left 91 participants to analyze from the Snowball population pool. The demographic information from these participants is located in Appendix B.1.

3.2.2.2 Stimuli

Eight United States Governor's voices were used in this study:

- Female Governors (high trust):
  - Michigan
  - South Dakota
- Female Governors (low trust):
  - Alabama

  - Iowa

- Male Governors (high trust):
  - Wisconsin
  - Vermont
- Male Governors (low trust):
  - Idaho
  - Arkansas

The samples were taken from the pilot study and the requirements for the sound bites can be found in Section 3.1.2.2. The samples selected from the pilot study were the two highest-rated and two lowest-rated female and male Governors. The criteria for selection came from a combination of the trust ratio for the COVID-19 trust question and the frequency responses for that question (i.e. the highest and lowest median and mode). Please note that although the pilot study sampled adequately for female Governors, the male Governors (excluding Indiana, Utah, and Virginia since the majority of the participants lived there, therefore limiting the number of responses for those Governors) were undersampled by half (only 7-9 raters for each male) of the necessary sample size suggested by Schild et al. (2020). This discrepancy between male and female ratings came about since there are 41 male Governors and 9 female Governors currently in office. Each participant had an even number of males and females to rate, which left more ratings per female Governor and less per male Governor. From this variability in the number of rankings for each Governor, the choice of male Governors came from the same selection criteria for the female Governors but focused on Governors who had almost all positive or negative responses. Using this method, bias was introduced during the selection process.

3.2.2.3 Procedure

Participants were given a link to a survey and completed it at their leisure with whatever audio equipment they use daily. The first question posed was about state residency. The answer to this question was used to curb potential bias towards their own Governor. That Governor was then removed from the list of potential sound bites. Participants were

then instructed to rate four randomly selected sound bites (two female and two males) on two questions:

1.) ”How trustworthy does the speaker sound?”
2.) ”How likely is it that this person will provide you with factual information about the COVID-19 pandemic?”

A 7-point Likert scale was used to answer each question with (1-very untrustworthy to 7-very trustworthy) for question 1 and (1- very unlikely to 7-very likely) for question 2. The full list of questions is provided in Appendix C.2.

An essential item to note is that this study deviates from the surrounding literature in that most other studies eliminate this confounding variable of semantic content by having the individuals say the same word or phrase (i.e. ”Hello”). Due to the nature of the COVID-19 pandemic, real sound bites were used from Governors during press conferences. The chosen sound bite was at the discretion of the author, meaning potential bias could have been injected into the study unknowingly. Rather than rating the Governors based on their prosody, participants may have instead rated these politicians on the validity and practicality of their reopening plans. All sound bites were chosen to include some amount of reopening plans (be it that the state was opening or remaining closed), but again, bias could have been introduced in this manner. To account for this bias, a semantic analysis was conducted using the SentimentAnnotator from Stanford’s CORE NLP library (v 4.0.0).[13] Each sentence was given a sentiment rating (-2: very negative to 2: very positive) and the average of this was analyzed to determine if there was any correlation between trust perceptions and the overall sentiment of their sound bite.

An individual’s prosody is also known to change dramatically from spontaneous speech to rehearsed narratives (Touati, 1993). Sound bites were selected based on rehearsed speech patterns (i.e. reading from a script) to eliminate the issues that arise from comparing different types of speech. However, while the Governors were reading their speeches, they may have adlibbed (injected spontaneous speech) and therefore changed their prosody

[13]https://stanfordnlp.github.io/CoreNLP/sentiment.html

throughout the clip. The author did not account for this directly, but instead, counted the number of ”ums” or ”uhs” in speech to indicate just how frequently a Governor may have injected their own thoughts and opinions into the rehearsed speech. The results from the relationship between these types of potential bias and trustworthiness measures are presented in the following sections.

3.2.3 Results

As in the Pilot Study, all sound bites were analyzed using Praat software (v 6.1.13) (Boersma and Van Heuven, 2001). The same process and settings were used. For the sake of brevity, those numbers have been omitted for the rest of this chapter. The reader is referred to Section 3.1.3 where the details are presented.

3.2.3.1 Trust Measures

To determine if any of the same correlations for articulation rate, phonation time, or standard deviation held true, Pearson's R, Spearman's R, and a two-tailed t-test with unequal variances were calculated. Pearson's R measures the linear relationships between two continuous variables whereas Spearman's R evaluates the monotonic relationship between ranked values. Since the data is Likert-type and not Likert-scale (i.e. taking a mean does not necessarily provide us useful data), we assume that our data may not have a linear relationship. In monotonic relationships, the variables change together, but not necessarily at a constant rate. Since we are not aware of the relationship at this time for any of the parameters and perceived trust, both metrics are used throughout this chapter. Most literature in this realm uses one-tailed t-tests because the authors are testing a specific hypothesis (i.e. ”Does higher mean F0 indicate higher trust in economic or mating contexts?”) and therefore know the direction they wish to test. A one-tail t-test allows researchers the ability to look more directly at a relationship. Since this study is testing correlations, not directions of correlations for the parameters we look at, we are not making any assumptions as to the direction of the relationship. A two-tail t-test is an appropriate way to explore these relationships.

		IA		MI		SD		AK		ID		VT		WI		AL	
		G	C	G	C	G	C	G	C	G	C	G	C	G	C	G	C
Frequency	1	1	2	0	0	1	2	11	12	2	4	2	1	2	4	6	9
	2	5	6	5	4	3	3	12	12	9	9	5	4	5	4	9	8
	3	13	6	7	3	3	1	23	24	14	13	6	6	3	4	15	19
	4	11	13	6	7	11	15	14	10	12	9	12	12	9	9	23	18
	5	28	26	14	12	25	21	24	18	18	26	21	18	36	27	20	17
	6	28	28	35	36	31	33	11	17	29	19	31	32	29	33	14	15
	7	8	13	22	27	20	19	1	3	6	10	14	18	9	12	6	7
Stats	3Q	6	6	6	7	6	6	5	5	6	6	6	6	6	6	5	5
	1Q	4	4	5	5	5	5	3	2.75	3	3	4	4.5	5	5	3	3
	IQR	2	2	1	2	1	1	2	2.25	3	3	2	1.5	1	1	2	2
	Median	5	5	6	6	6	6	4	3.5	5	5	5	6	5	5	4	4
	Mode	6	6	6	6	6	6	5	3	6	5	6	6	5	6	4	3

Table 3.3. COVID - Total: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the total participation pool. Abbreviations are as follows: IA = Iowa, MI = Michigan, SD = South Dakota, AK = Alaska, ID = Idaho, VT = Vermont, WI = Wisconsin, AL = Alabama. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

		IA		MI		SD		AK		ID		VT		WI		AL	
		G	C	G	C	G	C	G	C	G	C	G	C	G	C	G	C
Frequency	1	0	0	0	0	0	1	9	10	0	2	1	0	2	2	4	5
	2	3	3	1	2	0	1	7	7	5	5	3	3	2	1	5	4
	3	4	4	3	1	2	1	12	13	7	9	2	2	1	0	8	13
	4	6	6	2	1	3	5	8	5	9	4	5	4	3	8	16	11
	5	17	15	11	10	10	8	10	9	10	13	10	11	19	15	9	9
	6	13	15	16	19	18	18	1	3	10	8	16	17	15	16	3	3
	7	5	5	11	11	11	10	0	0	2	2	9	9	4	4	1	1
Stats	3Q	6	6	6.25	6.25	6.25	6	4	4.5	6	5	6	6	6	6	5	5
	1Q	4	4	5	5	5	5	2	2	3	3	5	5	5	5	3	3
	IQR	2	2	1.25	1.25	1.25	1	2	2.5	3	2	1	1	1	1	2	2
	Median	5	5	6	6	6	6	3	3	5	5	6	6	5	5	4	4
	Mode	5	5	6	6	6	6	3	3	6	5	6	6	5	6	4	3

Table 3.4. COVID - Snowball: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the Snowball participation pool. Abbreviations are as follows: IA = Iowa, MI = Michigan, SD = South Dakota, AK = Alaska, ID = Idaho, VT = Vermont, WI = Wisconsin, AL = Alabama. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

		IA		MI		SD		AK		ID		VT		WI		AL	
		G	C	G	C	G	C	G	C	G	C	G	C	G	C	G	C
Frequency	1	1	2	0	0	1	1	2	2	2	2	1	1	0	2	2	4
	2	2	3	4	2	3	2	5	5	4	4	2	1	3	3	4	4
	3	9	2	4	2	1	0	11	11	7	4	4	4	2	4	7	6
	4	5	7	4	6	8	10	6	5	3	5	7	8	6	1	7	7
	5	11	11	3	2	15	13	14	9	8	13	11	7	17	12	11	8
	6	15	13	19	17	13	15	10	14	19	11	15	15	14	17	11	12
	7	3	8	11	16	9	9	1	3	4	8	5	9	5	8	5	6
Stats	3Q	6	6	6	7	6	6	5	6	6	6	6	6	6	6	6	6
	1Q	3.25	4	4	5	4.25	4.25	3	3	3	4	4	4	5	5	3	3
	IQR	2.75	2	2	2	1.75	1.75	2	3	3	2	2	2	1	1	3	3
	Median	5	5	6	6	6	6	4	3.5	5	5	5	6	5	5	4	4
	Mode	6	6	6	6	5	6	5	6	6	5	6	6	5	6	5	6

Table 3.5. COVID - AMT: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the AMT participation pool. Abbreviations are as follows: IA = Iowa, MI = Michigan, SD = South Dakota, AK = Alaska, ID = Idaho, VT = Vermont, WI = Wisconsin, AL = Alabama. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

General Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.73	SD-41	194.51	112.80	18.32	4.62	5	0.2	34.18	36.98	-1	1
0.69	WI-49	140.24	120.30	18.37	3.72	4.31	0.232	33.66	38.96	-1	0
0.66	MI-22	197.97	229.60	37.61	3.36	4.47	0.224	30.41	40.48	-3	0
0.58	VT -45	156.38	170.40	22.89	3.82	4.65	0.215	36.55	44.47	0	0
0.48	IA -15	195.32	224	24.74	4.1	4.86	0.206	34.55	40.97	0	0
0.31	ID-12	133.19	159.40	30.41	2.72	4.35	0.23	23.46	37.47	-4	0
0.11	AL-1	195.69	217.50	32.13	4	5.12	0.195	28.7	36.71	1	3
-0.10	AK-2	144.71	237.10	31.44	4.13	5.2	0.192	31.73	39.98	-6	6
	Spearman	0.08	-0.62	-0.61	0.09	-0.62	0.52	0.50	0.12	0.23	-0.63
	Pearson	0.25	-0.60	-0.51	0.03	-0.57	0.55	0.43	0.19	0.42	-0.84
	P-Value	6.71E-07	1.69E-05	1.30E-05	1.93E-08	3.76E-13	7.82E-02	1.13E-07	6.59E-10	3.56E-02	3.29E-01

Table 3.6. All Politicians: This table provides the analysis of the vocal profile for each Governor and the Spearman and Pearson correlation coefficient for the "General Trust" condition. Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

COVID Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.76	Michigan	197.97	229.60	37.61	3.36	4.47	0.224	30.41	40.48	-3	0
0.71	South Dakota	194.51	112.80	18.32	4.62	5	0.2	34.18	36.98	-1	1
0.65	Wisconsin	140.24	120.30	18.37	3.72	4.31	0.232	33.66	38.96	-1	0
0.63	Vermont	156.38	170.40	22.89	3.82	4.65	0.215	36.55	44.47	0	0
0.56	Iowa	195.32	224	24.74	4.1	4.86	0.206	34.55	40.97	0	0
0.32	Idaho	133.19	159.40	30.41	2.72	4.35	0.23	23.46	37.47	-4	0
0.03	Alabama	195.69	217.50	32.13	4	5.12	0.195	28.7	36.71	1	3
-0.10	Alaska	144.71	237.10	31.44	4.13	5.2	0.192	31.73	39.98	-6	6

Spearman	-0.35	-0.59	-0.49	-0.62	-0.82	-0.58	-0.69	-0.57	-0.91	-0.86
Pearson	0.27	-0.48	-0.42	-0.02	-0.58	0.56	0.43	0.30	0.36	-0.85
P-Value	6.71E-07	1.69E-05	1.30E-05	1.42E-08	4.50E-13	8.31E-02	1.12E-07	6.19E-10	3.48E-02	3.36E-01

Table 3.7. All Politicians: This table provides the analysis of the vocal profile for each Governor and the Spearman and Pearson correlation coefficient for the "COVID Trust" condition. Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

3.2.4 Discussion

Table 3.3, Table 3.4, and Table 3.5 show the frequency distribution for the two population pools together, the Snowball participation pool, and the AMT participation pool respectively. Each of these tables shows the number of people who voted for each Governor in each category and some statistical parameters (interquartile range (IQR), median, and mode).

From the tables, we see that the Governors from Alabama, Idaho, and Arkansas were all ranked lowest. Iowa (who was the second-lowest female Governor ranked in the pilot study) is ranked in a similar fashion here. Although she was the second-lowest female Governor, she by no means was ranked ”low” and should not be considered in that light. The AMT participant pool voted for her slightly lower in general trust, but she was voted considerably higher overall for COVID (high-stress) trust. Stacked bar charts are provided with answers to the general trust question in Figures 3.1 - 3.3 and answers to the COVID trust question in Figures 3.4-3.6.

The stacked bar charts are ordered exactly the same for the lowest five Governors but for the upper three, they tend to change order, yet all remain in the top three. This order alteration should not be an indication that they are necessarily being rated different for either question, simply that they are so closely ranked in terms of ”high trust” that one participant who votes low for one question and high for another could skew the results dramatically. Upon taking a closer look at the difference between the AMT sample and the Snowball sample, it is evident that the AMT sample rankings were not spread as far as those in the Snowball sample. This represents that individuals in the AMT group tended to rate the Governors as higher overall compared to the Snowball sample. However, in both samples, it is evident that there is a clear distinction between the highest-rated and the lowest-rated Governor.

The Praat analysis and corresponding statistical tests for each metric are presented in Table 3.6 and Table 3.7 for answers to the general question and COVID question, respectively. For the general trust query, the range, standard deviation, and articulation rate were found to be statistically significant in all three tests. The highest correlation value for either the Spearman or Person correlation coefficient came from the ”Um” count, which checked for the number of times a person said ”um” or ”uh” during their sound bite. For the

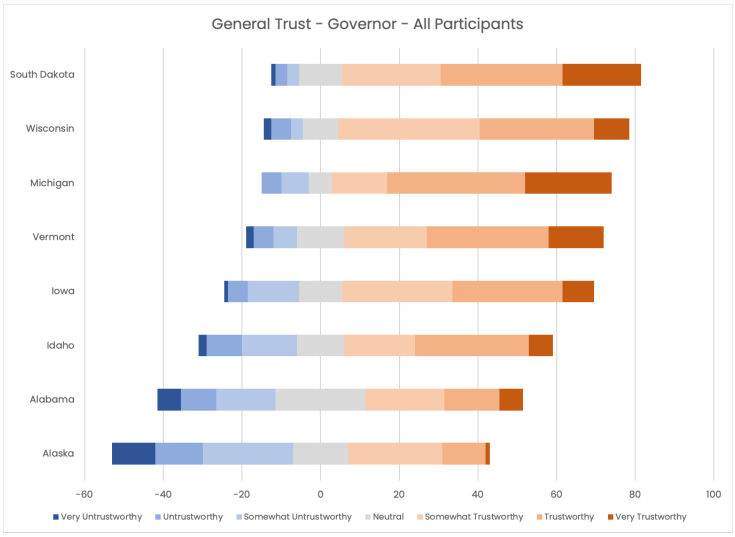

Figure 3.1. Frequency distribution of Governor's perceived trust for general trustworthiness by all participants. Ordered by highest ratio of positive to negative responses.

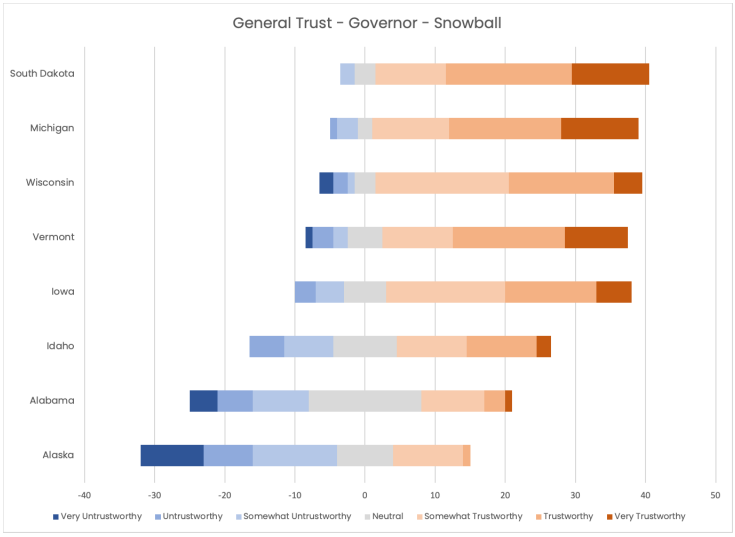

Figure 3.2. Frequency distribution of Governor's perceived trust for general trustworthiness by snowball participant pool. Ordered by highest ratio of positive to negative responses.

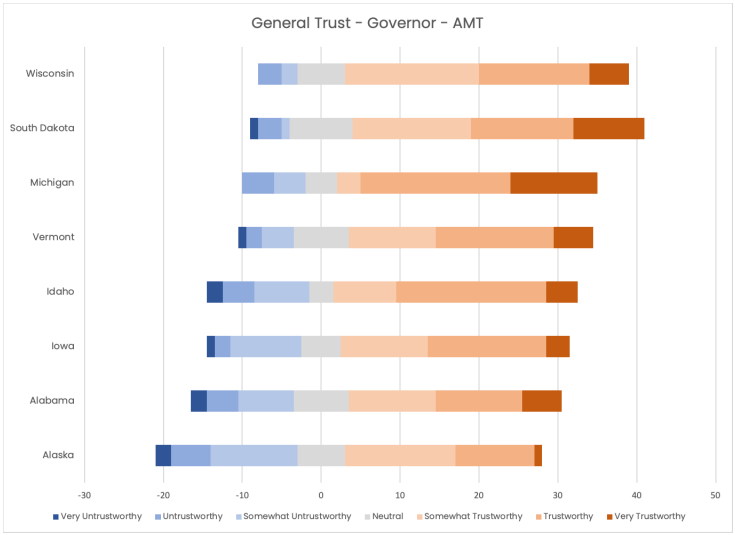

Figure 3.3. Frequency distribution of Governor's perceived trust for general trustworthiness by AMT participant pool. Ordered by highest ratio of positive to negative responses.

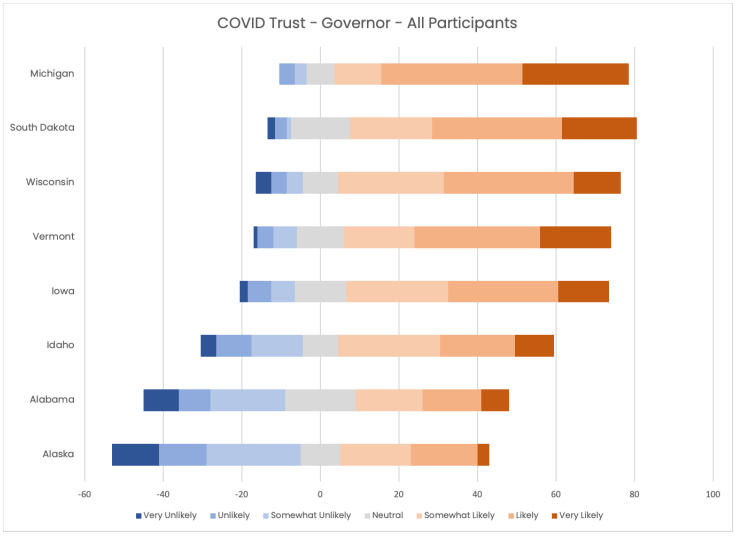

Figure 3.4. Frequency distribution of Governor's perceived trust for COVID trustworthiness by all participants. Ordered by highest ratio of positive to negative responses.

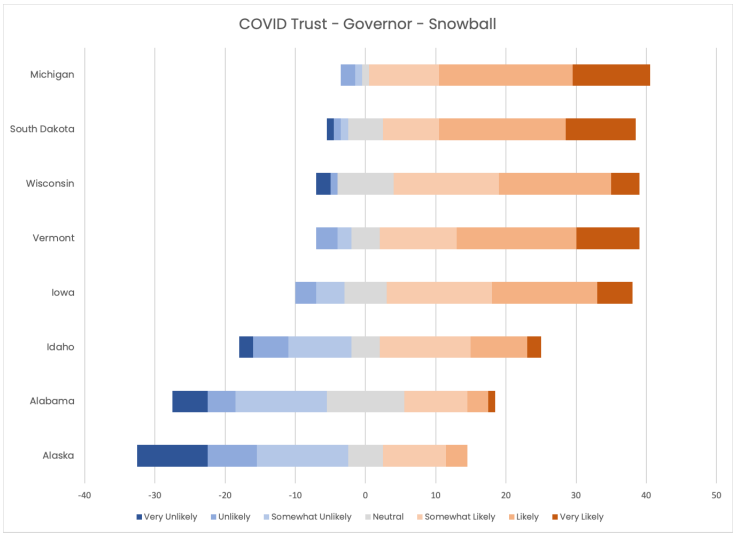

Figure 3.5. Frequency distribution of Governor's perceived trust for COVID trustworthiness by snowball participant pool. Ordered by highest ratio of positive to negative responses.

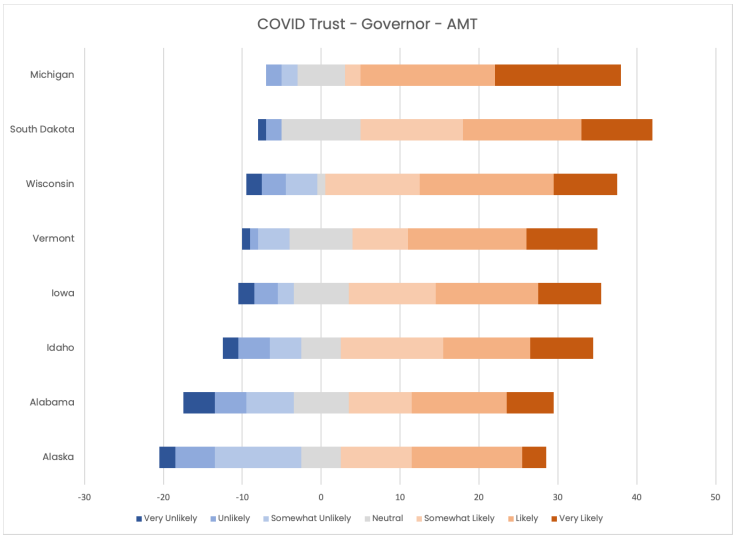

Figure 3.6. Frequency distribution of Governor's perceived trust for COVID trustworthiness by AMT participant pool. Ordered by highest ratio of positive to negative responses.

COVID question, the two parameters found to be significant across all three metrics were range and articulation rate, with the ”Um count” once again being the highest correlation value for both Spearman and Pearson. Phonation time and standard deviation were close to being statistically significant as well, but were slightly below the required value.

From these results, we can draw hypotheses regarding which prosodic elements correlate directly to Governor's' perceived trustworthiness. We were able to replicate the results found in the pilot study for how each Governor was ranked on the ”trust ratio.” It remains unclear what metric has the most influence on the independent variable (perceived trust) based on so few vocal profiles. Still unclear is the distinction between general trust and COVID trust (high-stress) rankings. The two values were nearly identical for all eight Governors. Experiment II builds on the results found here, but uses Mayors as the stimuli rather than Governors.

3.3 Experiment II: protests on Police Brutality & U.S. Mayors

3.3.1 Introduction

Protests and civil unrest ensued after the death of George Floyd on May 25, 2020. All across the United States, cities erupted in public demonstrations against police brutality, sometimes resulting in violence, and the National Guard was activated in at least 21 states. To provide context for how quickly these demonstrations materialized, a timeline is presented below[14]:

- 2020/05/25 - George Floyd, a 46-year-old African American man, died in Minneapolis, MN, after being handcuffed and pinned to the ground by a white police officer, Derek Chauvin. Bystanders recorded a video, and Mr. Floyd repeatedly said, "I can't breathe."
- 2020/05/26 - The video was shared widely on social media and prompted the protests in Minneapolis that evening. Medaria Arredondo, the Minneapolis police chief, fired all four men involved in the incident and called for an FBI investigation for the discrepancy between the police report and the video. Protests ensued that evening. Police vehicles were vandalized, and the precinct house where the four officers were assigned was a target for some demonstrators. Tear gas and rubber bullets were fired into the crowds.
- 2020/05/27 - Other cities began experiencing protests regarding Mr. Floyd, Breonna Taylor in Louisville, KY, and Ahmaud Arbery in Brunswick, Ga. One person was killed in both Chicago and St. Louis from the protests.
- 2020/05/28 - The National Guard was called up by the Minnesota Governor, Tim Walz.
- 2020/05/29 - Former officer Derek Chauvin was charged with third-degree murder and second-degree manslaughter. On Twitter, Mr. Trump called the protesters "thugs" and said, "When the looting starts, the shooting starts." Protests in Atlanta and New York that evening left destruction all around. Mr. Trump moved to an underground bunker.

[14] https://www.nytimes.com/article/george-floyd-protests-timeline.html

- 2020/05/30 - Minneapolis Mayor says peaceful protests have turned to ”domestic terrorism.”

- 2020/05/31 - More protests occurred, and the National Guard was deployed in more than two dozen states to assist the police.

- 2020/06/01 - George Floyd’s brother visits the site where he died. Two autopsies rule Mr. Floyd’s death a homicide but differed if other health complications contributed to his death. Trump threatens to deploy the military.

The ”Police and Criminal Evidence Act” mentioned previously, indirectly helped start these protests as a bystander was the one who recorded the incident. Without this video the police report would have given an entirely different narrative as to what happened. Although not all states require body cameras to be worn, citizens are now taking initiative to record all encounters with police to help validate their testimony.

The responses from both federal and state governments to these events were either appalling or adequate, depending on which news source captured the event. Once again the United States’ public was faced with a great deal of misinformation and misframing of the events. How does a constituent determine whom to trust for accurate portrayals of these events? Do they stay with the same rhetorical prosody from their favorite news station or are they now turning to different news sources because the broadcaster sounds informed? How can a person trust the information that is disseminated? Experiment II builds on Experiment I in that similar questions are posed to listeners at a time when civil unrest and data manipulation of current events are at extremes. Rather than concentrating on COVID-19, however, this Experiment focuses on Mayor’s responses to protests on police brutality.

3.3.2 Method

3.3.2.1 Participants

This study’s participants were split into two distinct participant groups: Snowball sampling and AMT. In both groups, participants were required to be 18 years or older and a current United States resident. These two groups were included to ensure that the results were from a more heterogeneous sample than either population independently. The

author compensated participants through AMT from his account. The compensation for the survey was based on the Qualtric's informed completion time for the survey (5 mins) and the federal minimum wage in the United States ($7.25). Therefore, each participant in the AMT population received $0.60 for completing the survey. The sample size was based off the same reasoning in Experiment I (see Section 3.2.2.1 for details).

104 participants in total accepted the HIT on AMT. To be considered for the HIT, the turker had to have one HIT approved and a HIT approval rating for all requesters' HITs higher than 95%. Once again, the reader is directed to Section 3.2.2.1 if they want to know more about why these requirements were selected. Out of the 104 participants, one entered the wrong survey code at the end, three responded to the attention checks incorrectly, and seven were under the 180 second time requirement. The time requirements were the same as Experiment I. After these requirements, 93 participants were left in the AMT population pool.

The Snowball sample consisted of participants from Twitter, Reddit, and LinkedIn. Overall, there were 118 survey responses from participants. 18 participants did not fully complete the survey and were eliminated from the pool. Four participants answered the attention checks incorrectly, one was not United States residents, two responses took under 180 seconds and five participants took over 3600 seconds. These requirements left 88 participants for the Snowball population pool. The demographic information from these participants is provided in Appendix B.2.

3.3.2.2 Stimuli

Eight United States Mayor's voices were used in this study:

- Los Angeles
- Phoenix
- Chicago
- New York
- Houston
- Las Vegas

- Oklahoma City
- San Antonio

These eight Mayors were determined based on the nature of the BLM movement. Police killing data gathered between Jan 2013 through Dec 2019 was analyzed and the top eight cities for the highest total number of police killings were chosen.[15]

Each sample follows the same requirements in section 3.1.2.2. These sound bites were gathered after June 1st, thus after the autopsy on Mr. Floyd showed that his death was a homicide and some protests around the country were starting to turn violent.

3.3.2.3 Procedure

Participants were given a link to a survey and completed it at their leisure with whatever audio equipment they use daily. The first question posed was about state residency. The answer to this question was used to curb potential bias towards their own Mayor. That Mayor was then removed from the list of potential sound bites. Participants were then instructed to rate four randomly selected sound bites on two questions:

1. ”How trustworthy does the speaker sound?”
2. ”How likely is it that this person will provide you with factual information regarding the protests on police brutality?”

A 7-point Likert scale was used to answer each question with (1-very untrustworthy to 7-very trustworthy) for 1 and (1- very unlikely to 7-very likely) for 2. Demographic questions were asked at the end with a heavy leaning on political thoughts and opinions. The full list of questions is provided in Appendix C.3.

3.3.3 Results

As in the Pilot Study, all sound bites were analyzed using Praat software (v 6.1.13) (Boersma and Van Heuven, 2001). The same process and settings were used. For the sake of brevity, those numbers have been omitted for the rest of this chapter. However, should

[15]https://mappingpoliceviolence.org/cities

the reader want to know more about these parameters, they are encouraged to refer to Section 3.1.3 where the details are presented.

3.3.3.1 Trust Measures

To determine if any of the same correlations for articulation rate, phonation time, or standard deviation held true, Pearson's R, Spearman's R, and a two-tailed t-test with unequal variances were calculated. The reasons for these choices are presented in Section 3.2.3.1.

		LA		CH		NY		HO		LV		OKC		SA		PH	
		G	P	G	P	G	P	G	P	G	P	G	P	G	P	G	P
Frequency	1	5	4	6	14	1	1	2	2	8	6	4	5	5	3	3	5
	2	3	6	17	6	6	6	2	2	6	14	9	15	5	5	11	7
	3	3	7	12	14	8	11	5	9	17	9	14	9	11	10	8	7
	4	14	13	10	12	12	11	9	11	15	14	11	11	5	11	8	5
	5	27	20	17	19	24	23	19	17	18	17	25	19	28	24	29	32
	6	19	22	26	24	31	23	38	31	17	24	26	26	26	24	31	31
	7	12	11	11	10	4	11	12	15	10	7	6	10	3	6	4	7
Stats	3Q	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	1Q	4	4	3	3	4	4	5	4	3	3	3	3	3.5	4	4	4.25
	IQR	2	2	3	3	2	2	1	2	3	3	3	3	2.5	2	2	1.75
	Median	5	5	5	5	5	5	6	6	4	5	5	5	5	5	5	5
	Mode	5	6	6	6	6	6	6	6	5	6	6	6	5	5	6	5

Table 3.8. Protests - Total: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the total participation pool. Abbreviations are as follows: LA = Los Angeles, CH = Chicago, NY = New York, HO = Houston, LV = Las Vegas, OKC = Oklahoma City, SA = San Antonio, PH = Phoenix. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

		LA		CH		NY		HO		LV		OKC		SA		PH	
		G	P	G	P	G	P	G	P	G	P	G	P	G	P	G	P
Frequency	1	4	3	5	11	1	1	1	1	4	4	1	3	3	2	0	1
	2	1	4	10	5	5	5	1	2	4	9	9	12	2	3	4	4
	3	2	3	8	8	4	7	4	8	12	7	10	5	7	5	6	7
	4	8	7	5	5	6	8	6	8	11	10	8	8	3	8	3	0
	5	18	13	7	6	9	7	10	8	7	10	10	6	15	14	12	14
	6	8	10	9	8	13	9	15	12	5	3	8	10	15	12	15	11
	7	2	3	3	4	1	2	7	5	3	3	1	3	0	1	3	6
Stats	3Q	5	6	5.5	5.5	6	6	6	6	5	5	5	6	6	6	6	6
	1Q	4	4	2	2	3.5	3	4	3.75	3	2	3	2	3	4	4	3
	IQR	1	2	3.5	3.5	2.5	3	2	2.25	2	3	2	4	3	2	2	3
	Median	5	5	4	3	5	4	5.5	5	4	4	4	4	5	5	5	5
	Mode	5	5	2	1	6	6	6	6	3	4	5	2	5	5	6	5

Table 3.9. Protests - Snowball: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the Snowball participation pool. Abbreviations are as follows: LA = Los Angeles, CH = Chicago, NY = New York, HO = Houston, LV = Las Vegas, OKC = Oklahoma City, SA = San Antonio, PH = Phoenix. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

		LA		CH		NY		HO		LV		OKC		SA		PH	
		G	P	G	P	G	P	G	P	G	P	G	P	G	P	G	P
Frequency	1	1	1	1	3	0	0	1	1	4	2	3	2	2	1	3	4
	2	2	2	7	1	1	1	1	0	2	5	0	3	3	2	7	3
	3	1	4	4	6	4	4	1	1	5	2	4	4	4	5	2	0
	4	6	6	5	7	6	3	3	3	4	4	3	3	2	3	5	5
	5	9	7	10	13	15	16	9	9	11	7	15	13	13	10	17	18
	6	11	12	17	16	18	14	23	19	12	21	18	16	11	12	16	20
	7	10	8	8	6	3	9	5	10	7	4	5	7	3	5	1	1
Stats	3Q	6.25	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
	1Q	4.75	4	4	4	5	5	5	5	4	4	5	4.75	4	4	4	5
	IQR	1.5	2	2	2	1	1	1	1	2	2	1	1.25	2	2	2	1
	Median	6	5.5	5	5	5	5	6	6	5	6	5	5	5	5	5	5
	Mode	6	6	6	6	6	5	6	6	6	6	6	6	5	6	5	6

Table 3.10. Protests - AMT: This table provides the frequency of each response including the interquartile range (IQR), third quartile (3Q), and first quartile (1Q) from the AMT participation pool. Abbreviations are as follows: LA = Los Angeles, CH = Chicago, NY = New York, HO = Houston, LV = Las Vegas, OKC = Oklahoma City, SA = San Antonio, PH = Phoenix. The cells highlighted in red indicate a low ranking, meaning a "low-trust perception" was recorded by participants.

General Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.69	Houston -5	122.90	116.80	22.90	3.03	4.45	0.225	29.41	43.22	0	0
0.57	LA - 1	121.16	123.10	18.08	3.1	4.54	0.22	25.8	37.71	-2	0
0.51	NY - 4	112.78	102.60	18.54	3.72	4.2	0.238	39.78	44.48	-5	0
0.45	Phoenix -2	187.78	182.3	25.06	4.27	5.05	0.198	37.64	44.47	-5	0
0.43	San Antonio -8	92.60	284.60	14.31	3.76	4.23	0.237	39.99	44.97	-3	10
0.32	OKC - 7	139.32	189.60	27.80	4.05	4.68	0.214	38.44	44.47	-3	8
0.19	Chicago -3	203.35	172.20	32.53	2.87	4.21	0.237	28.97	42.46	-5	0
0.15	Las Vegas -6	187.50	174.20	33.69	3.36	4.58	0.218	24.67	33.61	-2	0

Spearman	-0.87	-0.90	-0.94	-0.85	-0.76	-0.50	-0.77	-0.77	-0.03	-0.89
Pearson	-0.68	-0.39	-0.74	-0.02	0.00	0.01	0.17	0.36	0.39	-0.11
P-Value	2.17E-05	7.57E-05	2.80E-05	7.15E-08	4.30E-13	2.21E-02	2.01E-06	1.55E-08	8.38E-04	2.56E-01

Table 3.11. All Politicians: This table provides the analysis of the vocal profile for each Mayor and the Spearman and Pearson correlation coefficient for the "General Trust" condition. Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

Protest Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.57	Houston	122.90	116.80	22.90	3.03	4.45	0.225	29.41	43.22	0	0
0.54	Phoenix	187.78	182.3	25.06	4.27	5.05	0.198	37.64	44.47	-5	0
0.45	NY	112.78	102.60	18.54	3.72	4.2	0.238	39.78	44.48	-5	0
0.43	LA	121.16	123.10	18.08	3.1	4.54	0.22	25.8	37.71	-2	0
0.43	San Antonio	92.60	284.60	14.31	3.76	4.23	0.237	39.99	44.97	-3	10
0.27	OKC	139.32	189.60	27.80	4.05	4.68	0.214	38.44	44.47	-3	8
0.21	Las Vegas	187.50	174.20	33.69	3.36	4.58	0.218	24.67	33.61	-2	0
0.19	Chicago	203.35	172.20	32.53	2.87	4.21	0.237	28.97	42.46	-5	0

Spearman	-0.67	-0.82	-0.90	-0.68	-0.63	-0.61	-0.80	-0.78	-0.49	-0.92
Pearson	-0.51	-0.22	-0.69	0.24	0.23	-0.20	0.32	0.44	0.20	-0.11
P-Value	2.17E-05	7.56E-05	2.79E-05	1.31E-07	3.62E-12	1.52E-02	2.01E-06	1.57E-08	8.86E-04	2.51E-01

Table 3.12. All Politicians: This table provides the analysis of the vocal profile for each Mayor and the Spearman and Pearson correlation coefficient for the "Protest Trust" condition. Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

3.3.4 Discussion

Table 3.8, Table 3.9, and Table 3.10 show the frequency distribution for the two population pools together, the Snowball participation pool, and the AMT participation pool respectively. Each of these tables shows the number of people who voted for each Mayor in each category and some statistical parameters (IQR, median, and mode).

From the tables, we see that there was a slight difference in the way in which the Snowball sample and the AMT sample voted for perceived trustworthiness. For the Snowball sample, it was evident that Chicago, Las Vegas, and Oklahoma City were all voted significantly lower than the rest of the group. For AMT, however, the rankings differed for general trust and protest trust questions. The AMT and Snowball sample group ranked similarly when rating on the question of protests, but differed slightly when ranking on general trust. The most noticeable difference was how the populations voted for the Mayor of Oklahoma City. It should be noted here that the only sound bite the author could find for this Mayor was one in which the Mayor used more spontaneous speech than rehearsed speech. Some raters may have found the ”ums” and ”uhs” in his speech indicative of low general trust but high protest trust since the semantic content he provided demonstrated logical reasoning. Stacked bar charts are provided with answers to general trust questions in Figures 3.7 - 3.9 and answers to protests trust questions in Figures 3.10-3.12.

The stacked bar graphs help illustrate the differences in rankings and perhaps provide some amount of explanation as to why the rankings are different. For the AMT sample, the three female Governors had more very untrustworthy rankings than any other Mayor in the sample. This severely skewed the rankings for Mayor of Phoenix, who was ranked in the top three highest trust Mayors by the Snowball sample. The AMT sample, as in Experiment I, ranked every Mayor higher overall. Despite some discrepancies in the middle rankings, it is evident that Houston’s Mayor had ranked number one overall, and Chicago and Las Vegas were the two lowest-ranked Mayors.

Although it remains unclear why the two populations ranked some Mayors differently, it is evident that the demographics were drastically different between the two populations. After analysis, it was found that individuals aged 18-29, on average, ranked both Governors and Mayors as less trustworthy than any other age group. Heterosexuals, on average, ranked both Governors and Mayors as significantly more trustworthy than did

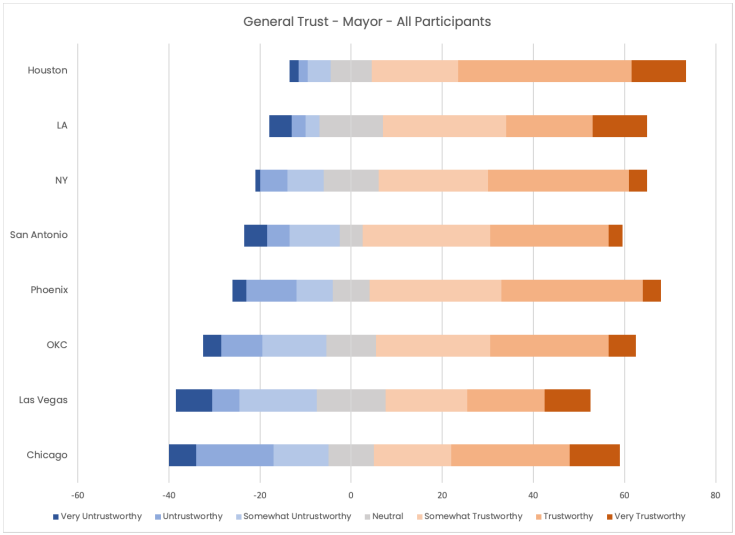

Figure 3.7. Frequency distribution of Mayor's perceived trust for general trustworthiness by all participants. Ordered by highest ratio of positive to negative responses.

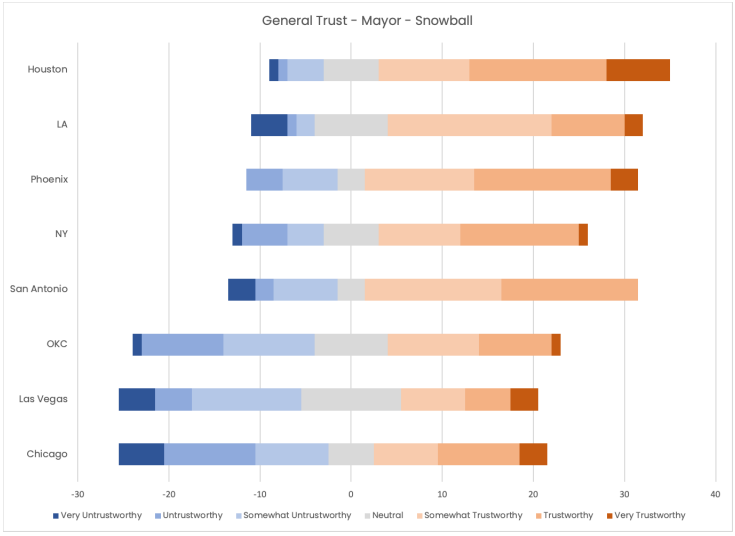

Figure 3.8. Frequency distribution of Mayor's perceived trust for general trustworthiness by snowball participant pool. Ordered by highest ratio of positive to negative responses.

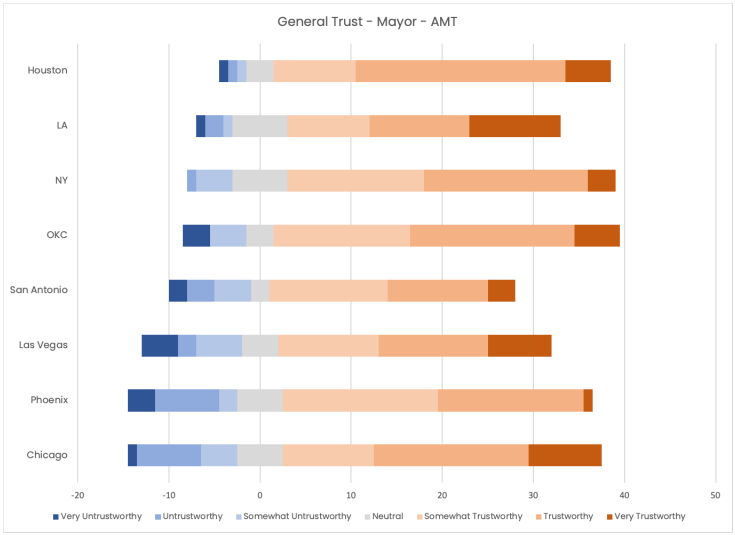

Figure 3.9. Frequency distribution of Mayor's perceived trust for general trustworthiness by AMT participant pool. Ordered by highest ratio of positive to negative responses.

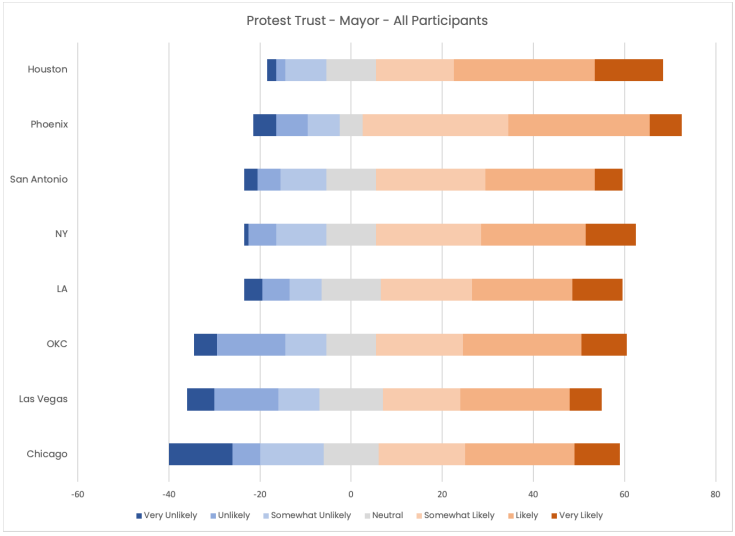

Figure 3.10. Frequency distribution of Mayor's perceived trust for Protest trustworthiness by all participants. Ordered by highest ratio of positive to negative responses.

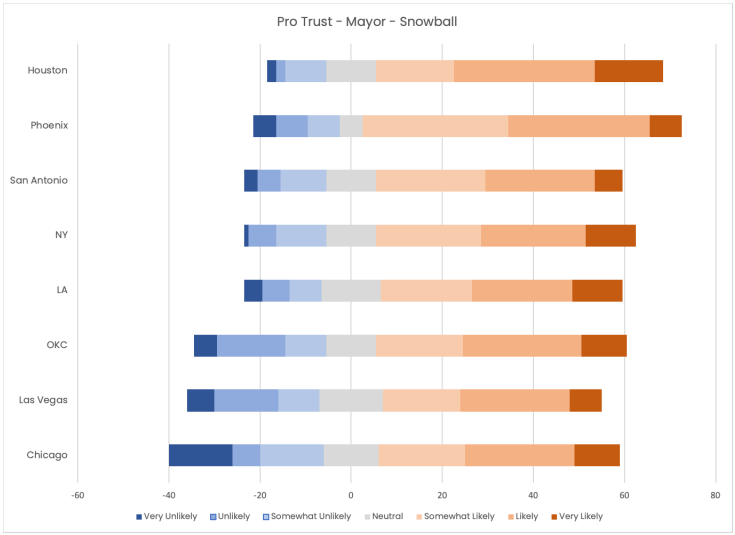

Figure 3.11. Frequency distribution of Mayor's perceived trust for Protest trustworthiness by snowball participant pool. Ordered by highest ratio of positive to negative responses.

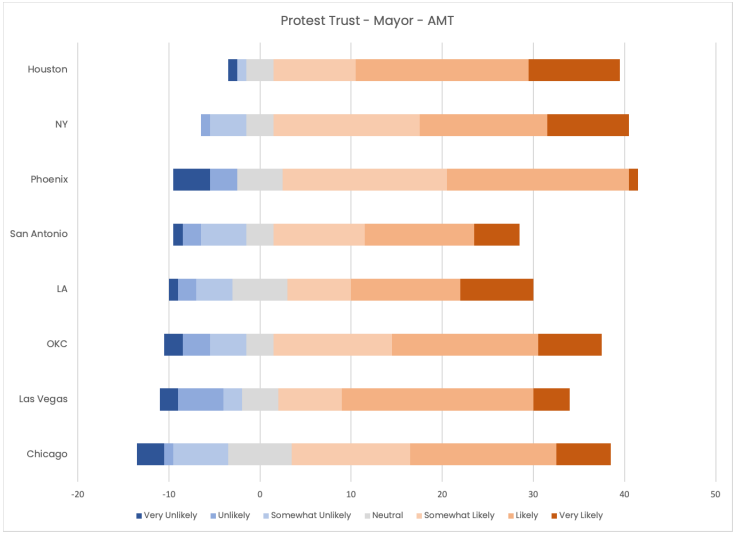

Figure 3.12. Frequency distribution of Mayor's perceived trust for Protest trustworthiness by AMT participant pool. Ordered by highest ratio of positive to negative responses.

participants who identify as non-heterosexual. Black participants, on average, ranked Governors and Mayors as high trust. Taken together, since the AMT population had far fewer participants aged 18-29, fewer participants that identified as non-heterosexual, and had a significantly higher black demographic, AMT rankings were higher overall for the Mayors as expected. Again, these demographics may not represent the real meaning behind why the two populations differed, but they do provide hypotheses to investigate further.

The Praat analysis and corresponding statistical tests for each metric are presented in Table 3.11 and Table 3.12 for answers to the general question and protest question, respectively. For the general trust query, the mean and standard deviation were statistically significant in all three tests. The highest correlation value for either the Spearman or Person correlation coefficient came from standard deviation. For the protests question, the two parameters found to be significant across all three metrics were again the mean and the standard deviation, with the ”Um” count being the highest correlation value for Spearman, slightly over standard deviation, but standard deviation being the highest for Pearson.

From these results, we can draw hypotheses regarding which prosodic elements correlate directly to Mayors' perceived trustworthiness. We were able to replicate results showing that people trust certain voices over others. We replicated that standard deviation is highly correlated to perceived trustworthiness, but no other parameters were found to be as highly correlated. It remains unclear as to what, if any, distinction exists between general trust and protest (high-stress) trust since the responses for each Mayor were nearly identical. In the next chapter, we will discuss how standard deviation plays a role in the broader scope of perceived trustworthiness and United States politicians.

CHAPTER 4

GENERAL DISCUSSION

In Section 3.1.4 (Pilot Study), we saw relationships indicating a strong correlation between perceived trust and standard deviation, articulation rate, and phonation time. Although we did not perform further statistical measures on these correlations (for lack of an appropriate sample size for male Governors), we saw that there seemed to be a trend in the data. In Section 3.2.4 (Experiment I), we found that the prosodic elements that were statistically significant across all three tests (Pearson, Spearman, and two-tailed t-test) were range, standard deviation, and articulation rate. The "Um" Count, although not statistically significant for the two-tailed t-test, was the highest correlated value for both the Pearson and Spearman correlation coefficients. In Section 3.3.4 (Experiment II), we found that the only prosodic elements that were statistically significant across all three tests (Pearson, Spearman, and two-tailed t-test) were the mean F0 and standard deviation. The "Um" count was the highest correlated value for the Spearman test but the standard deviation was the highest correlated value for the Pearson test.

Taken together, we see that standard deviation is the metric that best describes the differences in perceived trustworthiness for both the Governors and Mayors. Table 4.1 shows the prosodic elements of all the Governors and Mayors used in Experiment I and Experiment II ranked by their perceived trustworthiness ratio (the number of positive to negative responses) for general trust. Figure 4.2 shows the exact same elements but the trust ratio was determined by the answer to the situational trust question (i.e. either COVID or protests trust depending on whether the individual was a Governor or Mayor). Once again, the three tests (Spearman correlation, Pearson correlation, and two-tailed t-test with unequal variance in the data) were performed with the data collected from the two studies. For the general trust category, we found that there were no prosodic elements that

were statistically significant in all three categories. Rather, every one of them, besides the ”Um” count, was statistically significant for the t-tests. Articulation rate, average syllable duration, and duration were statistically significant for the Spearman coefficient, and the standard deviation and range were statistically significant for the Pearson coefficient. As for the situational trust answers, all prosodic elements were statistically significant for Spearman except the mean F0, and all prosodic elements were statistically significant for the t-tests except the ”Um” count. No prosodic element ended up being statistically significant for the Pearson coefficient, but standard deviation had the closest value to one that is statistically significant.

While many prosodic elements were statistically significant throughout these two experiments (excluding the pilot study), it remains a question of which parameters or combination of parameters best describe perceived trustworthiness. Throughout these three experiments (including pilot study) it is clear that standard deviation (and its close neighbor, range) plays some role in trust perception. To investigate this correlation in more depth, Fig 4.1 and Fig 4.2 are shown, depicting the relationship between the trust ratio and standard deviation for both general and situational trust. From these two graphs we see a clear negative relationship between the two variables with Governor Whitmer of Michigan as the outlier in the top right portion of each graph. If we remove this outlier, we will see that the relationship is even stronger than before and our ”almost statistically valid” correlations for standard deviation would most certainly be significant now. Governor Whitmer is considered an influential point, but is that the case because of the lack of vocal profiles with very high standard deviation or is this truly an outlier in the data? Another question to ask is why trust rankings tend to go down around 15Hz or so if the relationship is truly a negative correlation? If this were the case, the trust ratio should be higher while the standard deviation is low and keep decreasing as the standard deviation increases. Since we see a bend in the beginning, is there possibly another bend at the end?

General Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.73	SD-41	194.51	112.80	18.32	4.62	5	0.2	34.18	36.98	-1	1
0.69	Houston -5	122.90	116.80	22.90	3.03	4.45	0.225	29.41	43.22	0	0
0.69	WI-49	140.24	120.30	18.37	3.72	4.31	0.232	33.66	38.96	-1	0
0.66	MI-22	197.97	229.60	37.61	3.36	4.47	0.224	30.41	40.48	-3	0
0.58	VT -45	156.38	170.40	22.89	3.82	4.65	0.215	36.55	44.47	0	0
0.57	LA - 1	121.16	123.10	18.08	3.1	4.54	0.22	25.8	37.71	-2	0
0.51	NY - 4	112.78	102.60	18.54	3.72	4.2	0.238	39.78	44.48	-5	0
0.48	IA -15	195.32	224	24.74	4.1	4.86	0.206	34.55	40.97	0	0
0.45	Phoenix -2	187.78	182.3	25.06	4.27	5.05	0.198	37.64	44.47	-5	0
0.43	San Antonio -8	92.60	284.60	14.31	3.76	4.23	0.237	39.99	44.97	-3	10
0.32	OKC - 7	139.32	189.60	27.80	4.05	4.68	0.214	38.44	44.47	-3	8
0.31	ID-12	133.19	159.40	30.41	2.72	4.35	0.23	23.46	37.47	-4	0
0.19	Chicago -3	203.35	172.20	32.53	2.87	4.21	0.237	28.97	42.46	-5	0
0.15	Las Vegas -6	187.50	174.20	33.69	3.36	4.58	0.218	24.67	33.61	-2	0
0.11	AL-1	195.69	217.50	32.13	4	5.12	0.195	28.7	36.71	1	3
-0.10	AK-2	144.71	237.10	31.44	4.13	5.2	0.192	31.73	39.98	-6	6

Spearman	-0.14	0.28	0.40	-0.19	0.74	-1.00	0.04	0.80	-0.30	-0.07
Pearson	-0.14	-0.48	-0.56	0.02	-0.33	0.31	0.26	0.21	0.40	-0.38
P-Value	2.60E-11	1.03E-09	2.95E-10	8.65E-16	7.30E-26	4.02E-03	2.04E-13	1.32E-17	8.75E-05	1.27E-01

Table 4.1. All Politicians: This table provides the analysis of the vocal profile for each politician and the Spearman and Pearson correlation coefficient for the "General Trust" condition. Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

Situational Trust

Trust Ratio	State/City	Mean F0 (Hz)	Range	Std Dev	SR	AR	ASD	Phonation	dur (s)	Semantic	Um Count
0.76	Michigan	197.97	229.60	37.61	3.36	4.47	0.224	30.41	40.48	-3	0
0.71	South Dakota	194.51	112.80	18.32	4.62	5	0.2	34.18	36.98	-1	1
0.65	Wisconsin	140.24	120.30	18.37	3.72	4.31	0.232	33.66	38.96	-1	0
0.63	Vermont	156.38	170.40	22.89	3.82	4.65	0.215	36.55	44.47	0	0
0.57	Houston	122.90	116.80	22.90	3.03	4.45	0.225	29.41	43.22	0	0
0.56	Iowa	195.32	224	24.74	4.1	4.86	0.206	34.55	40.97	0	0
0.54	Phoenix	187.78	182.3	25.06	4.27	5.05	0.198	37.64	44.47	-5	0
0.45	NY	112.78	102.60	18.54	3.72	4.2	0.238	39.78	44.48	-5	0
0.43	LA	121.16	123.10	18.08	3.1	4.54	0.22	25.8	37.71	-2	0
0.43	San Antonio	92.60	284.60	14.31	3.76	4.23	0.237	39.99	44.97	-3	10
0.32	Idaho	133.19	159.40	30.41	2.72	4.35	0.23	23.46	37.47	-4	0
0.27	OKC	139.32	189.60	27.80	4.05	4.68	0.214	38.44	44.47	-3	8
0.21	Las Vegas	187.50	174.20	33.69	3.36	4.58	0.218	24.67	33.61	-2	0
0.19	Chicago	203.35	172.20	32.53	2.87	4.21	0.237	28.97	42.46	-5	0
0.03	Alabama	195.69	217.50	32.13	4	5.12	0.195	28.7	36.71	1	3
-0.10	Alaska	144.71	237.10	31.44	4.13	5.2	0.192	31.73	39.98	-6	6

Spearman	-0.41	-0.65	-0.70	-0.56	-0.72	-0.78	-0.78	-0.79	-0.64	-0.83
Pearson	0.01	-0.33	-0.43	0.08	-0.27	0.26	0.30	0.23	0.33	-0.41
P-Value	2.60E-11	1.03E-09	2.93E-10	7.18E-16	5.65E-26	5.52E-03	2.03E-13	1.30E-17	8.93E-05	1.25E-01

Table 4.2. All Politicians: This table provides the analysis of the vocal profile for each politician and the Spearman and Pearson correlation coefficient for the "Situational Trust" conditions (COVID and Protests). Abbreviations are as follows: SR = Speech Rate, AR = Articulation Rate (nsyll/phonation time), ASD = Average Syllable Duration (speaking time/nsyll). The cells highlighted in red indicate a statistically significant correlation between "Trust Ratio" and the metric being evaluated.

4.1 Additional Governors

To further investigate this query, vocal profiles from Governors in the pilot study were plot on the same figures to illustrate where they might lie on the trust ratio versus standard deviation relationship. Should the reader want to explore the prosodic elements of the Governors from the pilot study, Section 3.1.3 is where that information exists. Please note that this data from the pilot study is not necessarily exactly how these Governors would rank overall, but since Experiment I confirmed our rankings in the Pilot study, we will assume that the other rankings would remain similar and use this data to better postulate this relationship.

Figures 4.3 and 4.4 demonstrate what the relationship between trust ratio and standard deviation looks like with Governors from the Pilot Study and Experiment I and Mayors from Experiment II. From our initial hypothesis, we see the additional Governors map directly onto our relationships in both general and situational trust. Notice how Governor Cooper of North Carolina also has an extremely high standard deviation in his sound bite, and his trust ratio remains high. Governor Whitmer is no longer an anomaly, and the trust ratio arc from a standard deviation of about 10 to 35 Hz is even more prominent than it was previously. These four additional Governors help strengthen the argument that standard deviation plays an essential factor in trust perception; however, the two Governors with high trust and high standard deviation have yet to be determined as part of our population sample or merely influential points.

4.2 Prominent Historical Figures

To help validate our claims that these two Governors might be part of the population rather than influential points or outliers, four famous speeches in United States history were used as voice samples. Each sample was recorded and edited using the same hardware and software as previously mentioned (Section 3.1.2.2). The speech analysis was performed in the same manner as the other sound bites (Section 3.1.3). These speeches include:

- John F. Kennedy - ”Inauguration Speech”
- Martin Luther King Jr. - ”I have a dream”

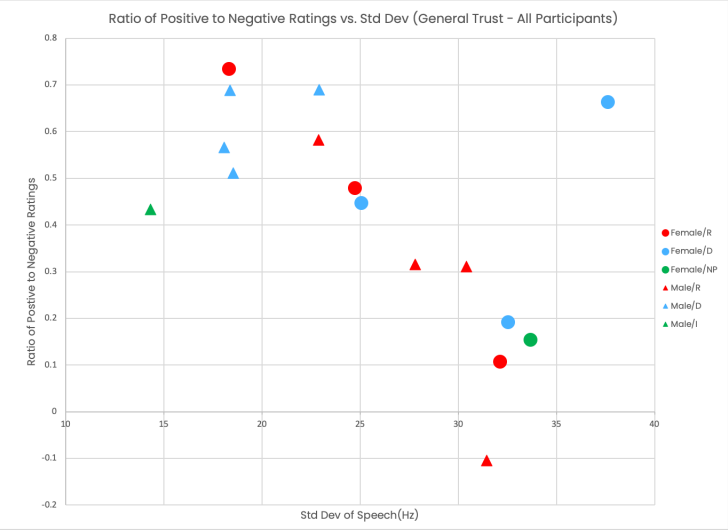

Figure 4.1. Ratio of Positive to Negative Rankings (General Trust Question) versus Standard Deviation in both Governors and Mayors.

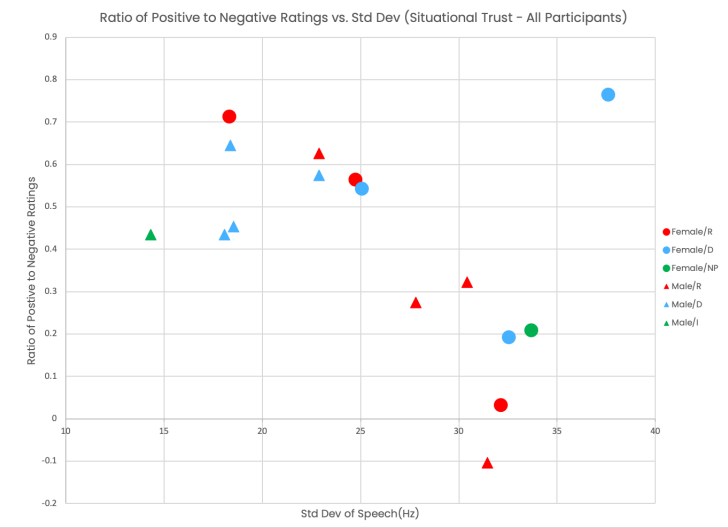

Figure 4.2. Ratio of Positive to Negative Rankings (Situational Trust Question) versus Standard Deviation in both Governors and Mayors.

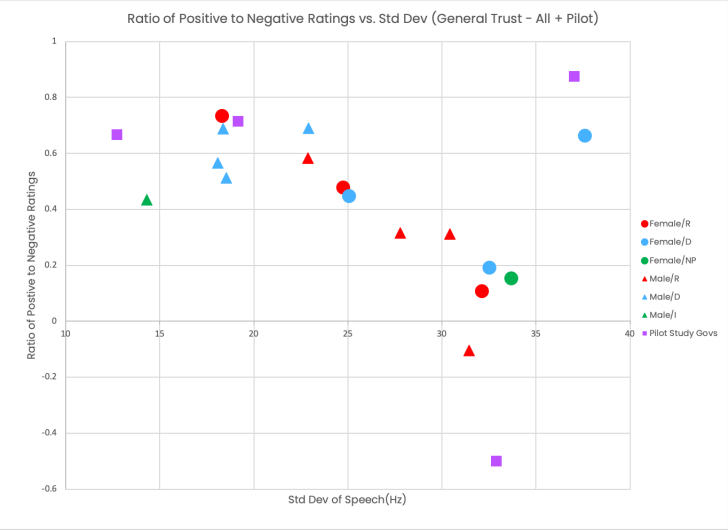

Figure 4.3. Ratio of Positive to Negative Rankings (General Trust Question) versus Standard Deviation in both Governors, Mayors and Pilot Study Governors.

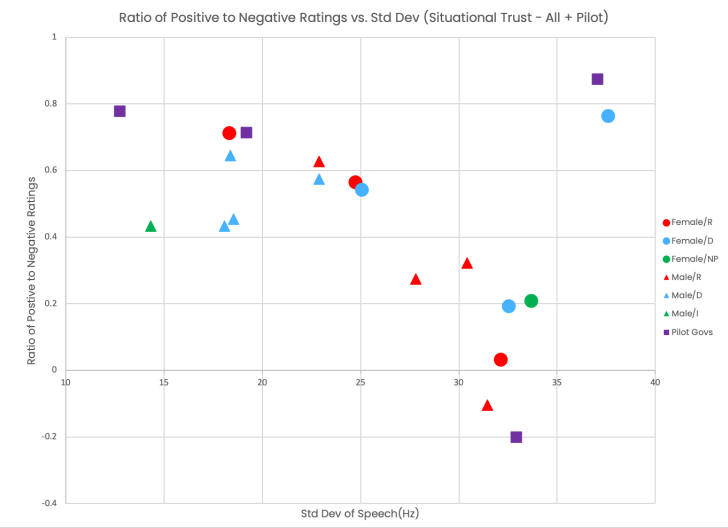

Figure 4.4. Ratio of Positive to Negative Rankings (Situational Trust Question) versus Standard Deviation in both Governors, Mayors, and Pilot Study Governors.

- Richard Nixon - ”Nixon’s First Watergate Speech”
- Bill Clinton - ”I did not have sexual relations with that woman”

No survey was needed for these prominent historical figures since we know how constituents perceived them during the time of their speech. MLK, JFK, and Bill Clinton were all perceived as high trust. MLK and JFK were considerably higher in trust than Bill Clinton; however, at the time of Clinton’s speech, many people still supported the claim that he did not have sexual relations with Ms. Lewinsky. On the other hand, Richard Nixon was perceived as low trust since his reputation was so poor that he had to resign from office. Since there was no way to measure the trust ratio for ”high” trust or ”low” trust without a survey, the first and third quartile of the survey data for trust rankings were used to denote low and high trust, respectively in both general and situational trust scenarios.

	Nixon	JFK	MLK	Clinton
notes	low	high	high	high
mean F0 (Hz)	113.08	196.44	250.81	116.96
range F0 (Hz)	232.7	280.3	291.6	115
std dv F0 (Hz)	29.06	35.95	40.75	19.21
speech rate (nsyll/dur)	2.9	4.16	3.8	3.34
articulation rate (nsyll / phonationtime)	4.81	4.46	5.21	3.34
ASD (speakingtime/nsyll)	0.208	0.224	0.192	0.299
phonation time	28.9	31.83	29.93	25.46
dur (s)	47.97	34.1	41.1	25.46

Table 4.3. Praat analysis of popular speeches in United States history.

Figures 4.5 and 4.6 denote the graphs with the Governors from both the Pilot study and Experiment I, Mayors from Experiment II, and the prominent historical figures mentioned previously. In both the General and situational trust scenarios, we see that the historical figures map directly onto our relationship. MLK and JFK have extremely high standard deviations and would have the highest trust ratio out of the four figures. From their addition to the graphs, it seems as if Governor Whitmer is less of an outlier now and is part of a trend where extremely high trust is often denoted by an extremely high standard deviation in the voice. Richard Nixon had a standard deviation closer to 30 Hz and fit along the curve with low trust, exactly how the Governors and Mayors had aligned. Clinton had a standard deviation of around 20 Hz and also fit into the cluster of highly rated Governors and Mayors around the same standard deviation.

These newly added figures beg us to answer the question: ”Why do we see a negative correlation between trust ratio and standard deviation up until a certain point?” I posit that we experience an ”uncanny valley,” as they denote it in Virtual Reality (VR) literature, in the modulation of pitch (i.e. standard deviation) from the speaker. If the reader is familiar with VR literature, they will be familiar with this nomenclature. There seems to be a curve up from 0Hz to 20Hz, dropping down at 30Hz, and rising back high at 35Hz. In VR, this ”uncanny valley” refers to when objects or items in the environment are real, but not real enough (Brenton et al., 2005). We usually see a positive correlation between the ”realism” of the environment and immersion scores. To better immerse the user into the VR, create a more realistic environment. However, there is a significant downward slope where the low fidelity environment or appearance is more immersive than the current one. Many papers investigate the reasoning for this in VR, but our talk remains in the auditory realm.

To better explain this stance, think about early voice assistants and how unnatural they sounded. Most of them were incredibly monotone (low standard deviation), and the pitch inflections were divergent from that of normal conversation. This point helps illustrate why speakers with lower standard deviations had a slightly lower perceived trust ranking. Something perhaps sounded a bit off, or the person made no expression through their voice. Although they were not rated ”lowest trust,” they also were not rated as the ”highest trust.” As we progress further towards higher standard deviations, we see this decline in trust ratio. I posit that we see this trend because people with this amount of standard

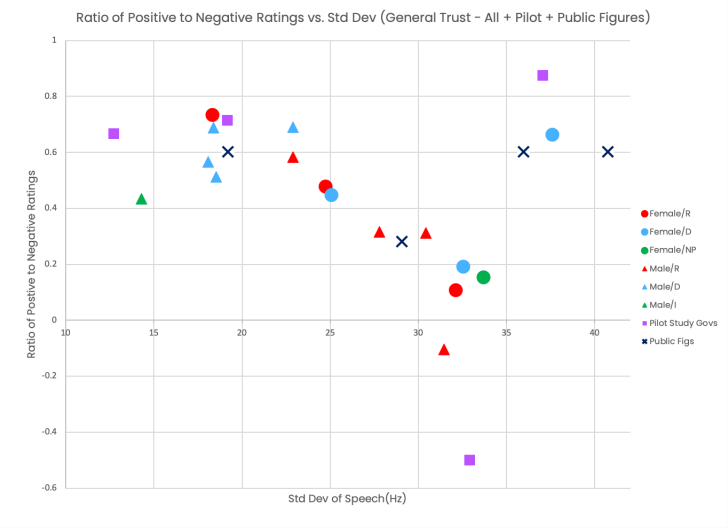

Figure 4.5. Ratio of Positive to Negative Rankings (General Trust Question) versus Standard Deviation in both Governors, Mayors, Pilot Study Governors, and Prominent Historical Figures.

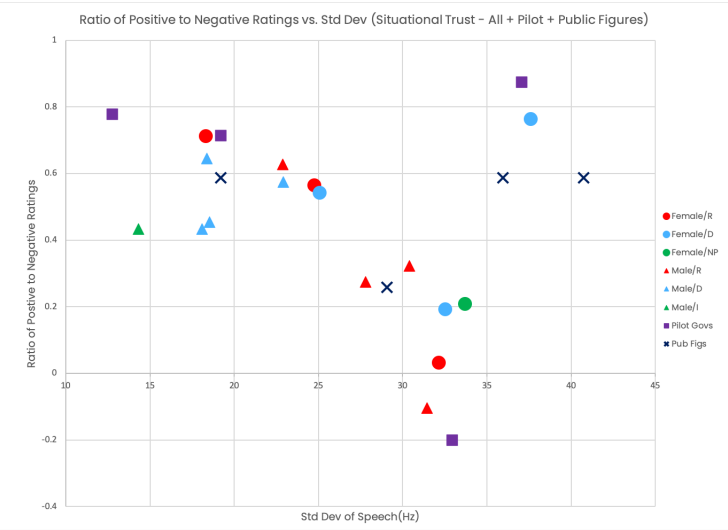

Figure 4.6. Ratio of Positive to Negative Rankings (Situational Trust Question) versus Standard Deviation in both Governors, Mayors, Pilot Study Governors, and Prominent Historical Figures.

deviation in their sound bite were often convincing, but not convincing enough. Rather than approaching the speech from that of inclusion, it may have sounded forceful and overall a little jarring. Recall a moment where a person's retelling of a story was less than convincing because they had dramatic pitch modulations such as "Well...anyway...we went to the store and you won't believe what happened!" These individuals display a pitch modulation that is higher than average speech, but not as high as the last category, which I call, "The Preacher Cluster." These orators use rhetorical prosody to not only convenience their listeners of their argument, but also in a way that resonates with a majority of people. Preachers often speak in this manner: high standard deviation, lower speech rate, and often a higher mean F0. The cluster of Governors and historical figures (no Mayors fell into this category), demonstrated in the previous figures, display these qualities in their voice sample. These qualities seem to emote a feeling of being talked to rather than being talked at. This inclusive nature may be the reason people rate these voices as high perceived trust. However, I posit that the curve will fall off around 47-50Hz because the pitch modulation would be too much for a listener. This kind of modulation would sound as if the person was singing rather than speaking and create another jarring reaction. For some individuals with Autism Spectrum Disorder, for example, their speech is described as "sing-song" and can lead to frustration when it comes to communication. On the other extreme, individuals with "monotonic" speech can also exhibit issues with communication, implying that low pitch modulation (standard deviation) creates a perception of low trust. Figure 4.7 and 4.8 are the same as Figures 4.5 and 4.6 with these added terms and a curve to indicate how this relationship might appear.

A final point that remains to be discussed is that of the difference between general trust and situational (COVID-19/protest) trust. There was little distinction between general trust and situational trust for all three experiments, leading the reader to believe that people do not make a significant distinction between the two. For this argument, I refer the reader to Schild et al. (2020). In trust literature there usually are three contexts explored: general, mating, and economic. Throughout this thesis I have argued that high-stress trust could be a different context to consider in future political psychology papers, such as the context of wartime presented in Tigue et al. (2012); Little et al. (2007). My results demonstrate that this is not entirely true and thus why I encourage you to look at the

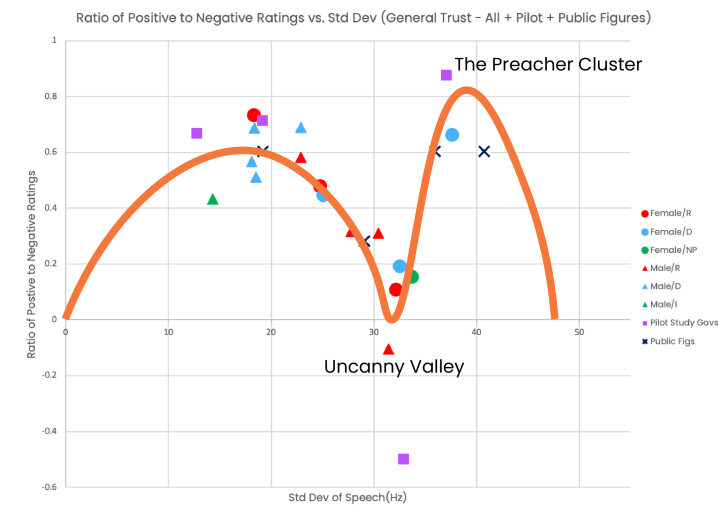

Figure 4.7. Ratio of Positive to Negative Rankings (General Trust Question) versus Standard Deviation in both Governors, Mayors, Pilot Study Governors, and Prominent Historical Figures.

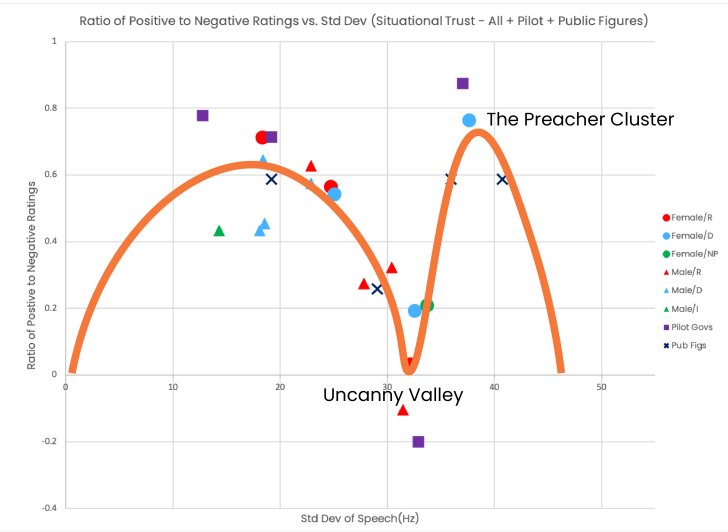

Figure 4.8. Ratio of Positive to Negative Rankings (Situational Trust Question) versus Standard Deviation in both Governors, Mayors, Pilot Study Governors, and Prominent Historical Figures.

results from Schild et al. (2020). The authors demonstrate that when it comes to prosodic elements, general trust and economic trust are nearly identical. It is instead the mating context that drastically differs from the other two. I argue that high-stress trust when it comes to prosodic elements in the voice and trust behaves identically to economic trust. Therefore, the main contribution of the thesis is not that high-stress trust behaves similarly to economic trust; rather, it is that the standard deviation does not behave linearly to perceived trust. In most studies in this field we believe there to be a linear positive or negative relationship to a prosodic element and trust. I demonstrate that this assumption may not be entirely accurate for every prosodic element. Future work should explore these relationships a bit more in-depth to determine the genuine interaction between each prosodic element and the perception of trust.

CHAPTER 5

CONCLUSION

The question posed initially in this thesis: ”Do prosodic elements of speech influence the perception of trustworthiness given to United States politicians?” has been challenged by the three experiments presented in Chapter 3. By using real-world samples, we had the additional challenge of ruling out semantic content and noisy backgrounds as confounding variables that may influence rater’s perceptions unknowingly. Although many prosodic elements of speech were found to be statistically significant in both Experiment I and Experiment II, the only prosodic element to show a strong correlation in both studies was the standard deviation or pitch modulation throughout the sound bite. When evaluating this metric with extra governors from the pilot study and four famous speeches from prominent historical figures, we saw that this correlation withstood these additional data points and we saw an even stronger trend emerge. The ”Uncanny Valley” was discussed as well as ”The Preacher Cluster” within the final two figures examining their presence in rehearsed narratives and the effect on perceived trust. Although standard deviation was found to hold true for these studies, future studies in political psychology and vocal trust literature should make note of this ”Uncanny Valley” in this specific prosodic element and examine relationships a bit closer when estimating vocal trust perception.

CHAPTER 6

REFERENCES

Alfarez Abdul-Rahman and Stephen Hailes. Supporting trust in virtual communities. In *Proceedings of the 33rd annual Hawaii international conference on system sciences*, pages 9–pp. IEEE, 2000.

David M Amodio and Chris D Frith. Meeting of minds: the medial frontal cortex and social cognition. *Nature reviews neuroscience*, 7(4):268–277, 2006.

Rindy C Anderson and Casey A Klofstad. Preference for leaders with masculine voices holds in the case of feminine leadership roles. *PloS one*, 7(12):e51216, 2012.

William Apple, Lynn A Streeter, and Robert M Krauss. Effects of pitch and speech rate on personal attributions. *Journal of personality and social psychology*, 37(5):715, 1979.

Robert Axelrod and William Donald Hamilton. The evolution of cooperation. *science*, 211 (4489):1390–1396, 1981.

Molly Babel, Grant McGuire, and Joseph King. Towards a more nuanced view of vocal attractiveness. *PloS one*, 9(2):e88616, 2014.

Irena Pavela Banai, Benjamin Banai, and Kosta Bovan. Vocal characteristics of presidential candidates can predict the outcome of actual elections. *Evolution and Human Behavior*, 38 (3):309–314, 2017.

Pat Barclay. Enhanced recognition of defectors depends on their rarity. *Cognition*, 107(3): 817–828, 2008.

Tamas Bereczkei, Anita Deak, Peter Papp, Gabor Perlaki, and Gergely Orsi. Neural correlates of machiavellian strategies in a social dilemma task. *Brain and cognition*, 82 (1):108–116, 2013.

Gordon W Blood, Blair Wallace Mahan, and Melvin Hyman. Judging personality and appearance from voice disorders. *Journal of communication disorders*, 12(1):63–67, 1979.

Paul Boersma. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings of the institute of phonetic sciences*, volume 17, pages 97–110. Amsterdam, 1993.

Paul Boersma and Vincent Van Heuven. Speak and unspeak with praat. *Glot International*, 5(9/10):341–347, 2001.

Aurélie Bonein and Daniel Serra. Gender pairing bias in trustworthiness. *The Journal of Socio-Economics*, 38(5):779–789, 2009.

Harry Brenton, Marco Gillies, Daniel Ballin, and David Chatting. The uncanny valley: does it exist. In *Proceedings of conference of human computer interaction, workshop on human animated character interaction*. Citeseer, 2005.

Bruce L Brown, William J Strong, and Alvin C Rencher. Acoustic determinants of perceptions of personality from speech. *International journal of the Sociology of Language*, 1975.

Judee K Burgoon, Thomas Birk, and Michael Pfau. Nonverbal behaviors, persuasion, and credibility. *Human communication research*, 17(1):140–169, 1990.

Ronald S Burt and Marc Knez. Kinds of third-party effects on trust. *Rationality and society*, 7(3):255–292, 1995.

John K Butler Jr. Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of management*, 17(3):643–663, 1991.

Danilo Bzdok, Robert Langner, Svenja Caspers, F Kurth, U Habel, Karl Zilles, A Laird, and Simon B Eickhoff. Ale meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Structure and Function*, 215(3-4):209–223, 2011.

Christiano Castelfranchi and Rino Falcone. *Trust theory: A socio-cognitive and computational model*, volume 18. John Wiley & Sons, 2010.

Daniel Cervone and Lawrence A Pervin. *Personality: Theory and research*. John Wiley & Sons, 2015.

Shelly Chaiken. Communicator physical attractiveness and persuasion. *Journal of Personality and social Psychology*, 37(8):1387, 1979.

Ananish Chaudhuri, Tirnud Paichayontvijit, and Lifeng Shen. Gender differences in trust and trustworthiness: Individuals, single sex and mixed sex groups. *Journal of Economic Psychology*, 34:181–194, 2013.

John Cook and Toby Wall. New work attitude measures of trust, organizational commitment and personal need non-fulfilment. *Journal of occupational psychology*, 53(1):39–52, 1980.

Elizabeth Couper-Kuhlen. *An introduction to English prosody*. Edward Arnold, 1986.

Alan Cruttenden et al. *Intonation*. Cambridge University Press, 1986.

Nivja H De Jong and Ton Wempe. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods*, 41(2):385–390, 2009.

LM DeBruine and BC Jones. Determining the number of raters for reliable mean ratings. *Open Science Framework. doi*, 10, 2018.

Timothy DeGroot and Stephan J Motowidlo. Why visual and vocal interview cues can affect interviewers' judgments and predict job performance. *Journal of Applied Psychology*, 84(6):986, 1999.

Timothy DeGroot, Federico Aime, Scott G Johnson, and Donald Kluemper. Does talking the talk help walking the walk? an examination of the effect of vocal attractiveness in leader effectiveness. *The Leadership Quarterly*, 22(4):680–689, 2011.

Julia R Dobrow and Calvin L Gidney. The good, the bad, and the foreign: The use of dialect in children's animated television. *The Annals of the American Academy of Political and Social Science*, 557(1):105–119, 1998.

Michael Doebeli, Christoph Hauert, and Timothy Killingback. The evolutionary origin of cooperators and defectors. *Science*, 306(5697):859–862, 2004.

Sarah Evans, Nick Neave, Delia Wakelin, and Colin Hamilton. The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4-5):783–788, 2008.

Rino Falcone and Cristiano Castelfranchi. Social trust: A cognitive approach. In *Trust and deception in virtual societies*, pages 55–90. Springer, 2001.

David R Feinberg, Benedict C Jones, Anthony C Little, D Michael Burt, and David I Perrett. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal behaviour*, 69(3):561–568, 2005.

Diego Gambetta et al. Can we trust trust. *Trust: Making and breaking cooperative relations*, 13:213–237, 2000.

Karl Grammer, Bernhard Fink, Anders P Møller, and Randy Thornhill. Darwinian aesthetics: sexual selection and the biology of beauty. *Biological Reviews*, 78(3):385–407, 2003.

Carlos Gussenhoven. Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science*, 8(2):425–434, 2016.

Carlos Gussenhoven et al. *The phonology of tone and intonation*. Cambridge University Press, 2004.

Eric Hehman, Sally Y Xie, Eugene K Ofosu, and Gabe Nespoli. Assessing the point at which averages are stable: A tool illustrated in the context of person perception. 2018.

Caroline G Henton and R Anthony Bladon. Breathiness in normal female speech: Inefficiency versus desirability. *Language & Communication*, 1985.

Julia Bell Hirschberg and Andrew Rosenberg. Acoustic/prosodic and lexical correlates of charismatic speech. 2005.

Jill House. The role of prosody in constraining context selection: a procedural approach. *Nouveaux cahiers de linguistique française*, 28:369–383, 2007.

Susan M Hughes, Franco Dispenza, and Gordon G Gallup Jr. Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25 (5):295–304, 2004.

Xiaoming Jiang and Marc D Pell. The sound of confidence and doubt. *Speech Communication*, 88:106–126, 2017.

Cynthia Johnson-George and Walter C Swap. Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of personality and social psychology*, 43(6):1306, 1982.

Gareth R Jones and Jennifer M George. The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of management review*, 23(3):531–546, 1998.

Christin Kirchhübel and David M Howard. Detecting suspicious behaviour using speech: Acoustic correlates of deceptive speech–an exploratory investigation. *Applied ergonomics*, 44(5):694–702, 2013.

Dennis H Klatt and Laura C Klatt. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *the Journal of the Acoustical Society of America*, 87(2):820–857, 1990.

C Klofstad. Candidate voice pitch influences voters. In *Midwest Political Science Association Annual Conference. Chicago, IL, USA*, 2013.

Casey A Klofstad. Candidate voice pitch influences election outcomes. *Political Psychology*, 37(5):725–738, 2016.

Casey A Klofstad, Rindy C Anderson, and Susan Peters. Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738):2698–2704, 2012.

Tobias L Kordsmeyer, John Hunt, David A Puts, Julia Ostner, and Lars Penke. The relative importance of intra-and intersexual selection on human male sexually dimorphic traits. *Evolution and Human Behavior*, 39(4):424–436, 2018.

Timothy R Koscik and Daniel Tranel. The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia*, 49(4):602–611, 2011.

Michael Kosfeld. Trust in the brain: Neurobiological determinants of human social behaviour. *EMBO reports*, 8(S1):S44–S47, 2007.

Robert V Krejcie and Daryle W Morgan. Determining sample size for research activities. *Educational and psychological measurement*, 30(3):607–610, 1970.

Eva Krumhuber, Antony SR Manstead, Darren Cosker, Dave Marshall, Paul L Rosin, and Arvid Kappas. Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, 7(4):730, 2007.

D Robert Ladd. *Intonational phonology*. Cambridge University Press, 2008.

Lasse Laustsen, Michael Bang Petersen, and Casey A Klofstad. Vote choice, ideology, and social dominance orientation influence preferences for lower pitched voices in political candidates. *Evolutionary Psychology*, 13(3):1474704915600576, 2015.

John DM Laver. Voice quality and indexical information. *British Journal of Disorders of Communication*, 3(1):43–54, 1968.

Shiri Lev-Ari and Boaz Keysar. Why don't we believe non-native speakers? the influence of accent on credibility. *Journal of experimental social psychology*, 46(6):1093–1096, 2010.

J David Lewis and Andrew Weigert. Trust as a social reality. *Social forces*, 63(4):967–985, 1985.

Rosina Lippi, Sara Donati, Rosina Lippi-Green, and Rosina Donati. *English with an accent: Language, ideology, and discrimination in the United States*. Psychology Press, 1997.

Anthony C Little, Robert P Burriss, Benedict C Jones, and S Craig Roberts. Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1):18–27, 2007.

Sara Mack and Benjamin Munson. The influence of/s/quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp'stereotype. *Journal of Phonetics*, 40(1):198–212, 2012.

Roger C Mayer, James H Davis, and F David Schoorman. An integrative model of organizational trust. *Academy of management review*, 20(3):709–734, 1995.

Phil McAleer, Alexander Todorov, and Pascal Belin. How do you say 'hello'? personality impressions from brief novel voices. *PloS one*, 9(3):e90779, 2014.

Daniel J McAllister. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal*, 38(1):24–59, 1995.

Robert R McCrae. The five-factor model of personality traits: Consensus and controversy. *The Cambridge handbook of personality psychology*, pages 148–161, 2009.

D Harrison McKnight, Vivek Choudhury, and Charles Kacmar. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *The journal of strategic information systems*, 11(3-4):297–323, 2002.

Barbara Misztal. *Trust in modern societies: The search for the bases of social order*. John Wiley & Sons, 1996.

Carlos David Navarrete, Melissa M McDonald, Michael L Mott, Joseph Cesario, and Robert Sapolsky. Fertility and race perception predict voter preference for barack obama. *Evolution and Human Behavior*, 31(6):394–399, 2010.

Daniel Nettle and Bethany Liddle. Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality: Published for the European Association of Personality Psychology*, 22(4):323–335, 2008.

Richard E Nisbett and Timothy D Wilson. The halo effect: evidence for unconscious alteration of judgments. *Journal of personality and social psychology*, 35(4):250, 1977.

Jillian JM O'Connor and Pat Barclay. The influence of voice pitch on perceptions of trustworthiness across social contexts. *Evolution and Human Behavior*, 38(4):506–512, 2017.

Jillian JM O'Connor, Paul J Fraccaro, Katarzyna Pisanski, Cara C Tigue, Timothy J O'Donnell, and David R Feinberg. Social dialect and men's voice pitch influence women's mate preferences. *Evolution and Human Behavior*, 35(5):368–375, 2014.

Craig Packer. Reciprocal altruism in papio anubis. *Nature*, 265(5593):441–443, 1977.

Dennis R Preston. *Handbook of perceptual dialectology*, volume 1. John Benjamins Publishing, 1999.

David Andrew Puts, Steven JC Gaulin, and Katherine Verdolini. Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and human behavior*, 27(4):283–296, 2006.

Drew Rendall, John R Vokey, and Christie Nemeth. Lifting the curtain on the wizard of oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5):1208, 2007.

Betty Repacholi, Virginia Slaughter, Michelle Pritchard, and Vicki Gibbs. Theory of mind, machiavellianism, and social functioning in childhood. 2003.

Martin Reuter, Christian Montag, Stephen Altmann, Fabian Bendlow, Christian Elger, Peter Kirsch, B Weber, and Armin Falk. Genetically determined differences in human trust behavior: The role of the oxytocin receptor gene. In *NeuroPsychoEconomics conference proceedings*, volume 21, 2009.

René Riedl and Andrija Javor. The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. *Journal of Neuroscience, Psychology, and Economics*, 5(2):63, 2012.

Jonathan Robinson, Cheskie Rosenzweig, Aaron J Moss, and Leib Litman. Tapped out or barely tapped? recommendations for how to harness the vast and largely unused potential of the mechanical turk participant pool. *PloS one*, 14(12):e0226394, 2019.

Milton J Rosenberg, Carl I Hovland, William J McGuire, Robert P Abelson, and Jack W Brehm. Attitude organization and change: An analysis of consistency among attitude components.(yales studies in attitude and communication.), vol. iii. 1960.

Bo Rothstein and Eric M Uslaner. All for all: Equality, corruption, and social trust. *World Pol.*, 58:41, 2005.

Julian B Rotter. A new scale for the measurement of interpersonal trust. *Journal of personality*, 1967.

Nicholas O Rule, Michael L Slepian, and Nalini Ambady. A memory advantage for untrustworthy faces. *Cognition*, 125(2):207–218, 2012.

Christoph Schild, Julia Stern, and Ingo Zettler. Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, 31(1):164–175, 2020.

Aaron Sell, Gregory A Bryant, Leda Cosmides, John Tooby, Daniel Sznycer, Christopher Von Rueden, Andre Krauss, and Michael Gurven. Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B: Biological Sciences*, 277 (1699):3509–3518, 2010.

Chilin Shih and Greg Kochanski. What is prosody?, 2002. URL `http://www.cs.columbia.edu/ julia/cs4706/chilin.htm`.

Robert Slonim and Pablo Guillen. Gender selection discrimination: Evidence from a trust game. *Journal of Economic Behavior & Organization*, 76(2):385–405, 2010.

Melissa K Surawski and Elizabeth P Ossoff. The effects of physical and vocal attractiveness on impression formation of politicians. *Current Psychology*, 25(1):15–27, 2006.

Helen Tager-Flusberg and Kate Sullivan. A componential view of theory of mind: evidence from williams syndrome. *Cognition*, 76(1):59–90, 2000.

Cara C Tigue, Diana J Borak, Jillian JM O'Connor, Charles Schandl, and David R Feinberg. Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3):210–216, 2012.

Alexander Todorov, Chris P Said, Andrew D Engell, and Nikolaas N Oosterhof. Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences*, 12(12): 455–460, 2008.

Paul Touati. Temporal profiles and tonal configurations in french political speech. *Lund Working Papers in Linguistics*, 38:205–219, 1991.

Paul Touati. Prosodic aspects of political rhetoric. In *ESCA workshop on prosody*, 1993.

Robert L Trivers. The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1): 35–57, 1971.

Peter Trudgill et al. *Sociolinguistics: An introduction to language and society*. Penguin UK, 2000.

Gina Villar, Joanne Arciuli, and Helen Paterson. Vocal pitch production during lying: Beliefs about deception matter. *Psychiatry, Psychology and Law*, 20(1):123–132, 2013.

Yao Wang and Julita Vassileva. Trust and reputation model in peer-to-peer networks. In *Proceedings Third International Conference on Peer-to-Peer Computing (P2P2003)*, pages 150–157. IEEE, 2003.

Stuart A West, Claire El Mouden, and Andy Gardner. Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and human behavior*, 32(4):231–262, 2011.

Janine Willis and Alexander Todorov. First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological science*, 17(7):592–598, 2006.

Yi Xu, Albert Lee, Wing-Li Wu, Xuan Liu, and Peter Birkholz. Human vocal attractiveness as signaled by body size projection. *PloS one*, 8(4):e62397, 2013.

Toshio Yamagishi and Midori Yamagishi. Trust and commitment in the united states and japan. *Motivation and emotion*, 18(2):129–166, 1994.

Toshio Yamagishi, Shigehito Tanida, Rie Mashima, Eri Shimoma, and Satoshi Kanazawa. You can judge a book by its cover: Evidence that cheaters may look different from cooperators. *Evolution and Human Behavior*, 24(4):290–301, 2003.

Paul J Zak, Robert Kurzban, and William T Matzner. Oxytocin is associated with human trustworthiness. *Hormones and behavior*, 48(5):522–527, 2005.

Miron Zuckerman and Robert E Driver. What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, 13(2):67–82, 1989.

Miron Zuckerman and Kunitate Miyake. The attractive voice: What makes it so? *Journal of nonverbal behavior*, 17(2):119–135, 1993.

APPENDIX A

TRANSCRIPTS

A.1 Governors

A.1.1 Gov 1

"*uh* stay at home order stays in effect through April 30th.

Last week I talked to you 'bout the *uh* shift in focus from our people's safety and health to include also the *uh* economic health of our state.

In our mind those two have been closely interwoven since they very beginning.

As most people can appreciate this requires a strategic approach as well as a team effort to strike the proper balance between keeping our people safe and healthy and also addressing our state's overall economic health."

Um Count: 3

Sentiment Analysis: 1

A.1.2 Gov 2

"But we have been careful.

um we've been careful from the beginning with your help, again.

And so *um* we're gonna talk about *um* *um* again what *um* what we're looking at reopening.

And *um* it's gonna happen pretty quickly.

But before we go there I want to answer a couple questions from Friday that some of the reporters had asked us.

Sir Aaron McGrowery from the News Meyer says "Will you use federal relief money to provide hazard pay to healthcare workers?"

And of course we're, I mean, we're deeply appreciative of what everyone is doing on the front lines including healthcare workers."

Um Count: 6

Sentiment Analysis: -6

A.1.3 Gov 12

"I know this is hard on everyone.

I have spoken with many workers, business owners, who have shared their stories of hardship, concern, and outright anger.

I've heard from many, many people who are upset that our current circumstances prevent them from practicing their faith the way they want.

Parents are frustrated.

Children long for their teachers and classmates.

Wedding plans, travel plans, incomes, and personal spending all have changed."

Um Count: 0

Sentiment Analysis: -4

A.1.4 Gov 15

"And as you've seen daily through the RMCC reports, our ICU beds, vents are well managed and in good supply.

The coordination and the collaboration between providers, systems, and the state gives us the confidence that we can effectively care for COVID-19 patients while also providing procedures to improve health and quality of life for others.

On Monday, each hospital outpatient surgery center or clinic that determines they can safely do so, may begin.

They may begin rescheduling patients and resume surgeries and procedures according to their own schedules."

Um Count: 0

Sentiment Analysis: 0

A.1.5 Gov 22

"It's good, but we must keep it up.

The order I signed today requires that everyone wear a covering over their nose and mouth like a homemade mask or scarf or bandana or handkerchief when they're in an enclosed public space.

To be clear, when you go to the grocery store, or to the pharmacy, or to any store that is open during this time, you need to wear a face mask.

If you're in an outdoor area, this order doesn't require that you wear one, but you should consider it anyway.

And regardless, you need to observe the six feet radius of safety so that you are protected from spread."

Um Count: 0

Sentiment Analysis: -3

A.1.6 Gov 41

"We're being aggressive and making sure that people are taking the personal responsibility to be disciplined in their actions each and every day and how they conduct their way of life.

But that's why I'm continuing to remind you of your responsibility.

Um..make sure that you are do...taking these actions, not just for your own personal health, but for those of your family and your community members.

I also want to remind you to download the CARE-19 app.

We've had 14,000 people download that app so far.

I need many, many more of you to do that.

Remember this will help us make sure that we're protecting as many people from exposure to the virus in the coming days."

Um Count: 1

Sentiment Analysis: -1

A.1.7 Gov 45

”Now before I get into the steps we’re taking today I want to remind everyone of the principles we set that guide any re-openings.

First, we’ll continue to keep our eyes on the data and make sure we pay attention to what’s happening and that we continue to move in the right direction.

Second, we will make sure our health care system remains ready so we’re prepared to fight outbreaks as they happen and we have the capacity to treat COVID patients.

Third, we will continue to work with our state lab, hospitals, and commercial labs and make testing more available with a focus on proactive testing and followed up by building upon our strong contact tracing program.”

Um Count: 0

Sentiment Analysis: 0

A.1.8 Gov 49

”Good afternoon and welcome back, after the weekend.

As you know last week I extended our safter at home order.

This is not an easy decision to make, but as I’ve said all along, the health and safety of the people of our state comes first.

But at the same time, we also know that we have to do everything we can to reopen our state as soon as we can safely and responsibly can while recognizing that we will not get back to normal until we have a vaccine.

No one wants to reopen our economy as much as I do, but folks, like I’ve said before, it has to be more like turning a dial than flipping a light switch.”

Um Count: 0

Sentiment Analysis: -1

A.2 Mayors

A.2.1 Mayor 1

"I left the protest in the street, after taking a knee, after praying, and after addressing the crowd.

And I joined rev...reverend K. W. Tullos, who's the president of the Baptists Minister's Conference.

My old friends, Pastor William Smart, Bishop Juan Carlos Mendez, and Pastor Michael Fisher Junior who's here with me tonight and many others.

And before anybody who's friends with them say anything to them, they gave me a good, hard time.

They were demanding about justice as everyone in America should be right now."

Um Count: 0

Sentiment Analysis: -2

A.2.2 Mayor 2

"However, also attending the protests, was a smaller, but dangerous, group of individuals who thought that the way to honor a stolen life was to break windows, angrily confront reporters, and vandalize both government and private buildings.

This damage was not acceptable.

We know these individuals do not represent people who...do not represent most people who took to our streets.

We also will not allow violent actions of these individuals to overshadow the peaceful protests being practiced by other residents.

We cannot let a small group, propensity for violence and chaos, direct our attention away from the very real conversation that needs to be had about race and policing in America.

We will not stand for this violence and destruction."

Um Count: 0

Sentiment Analysis: -5

A.2.3 Mayor 3

"I watched as protesters hurled not just words, or projectiles, at our police department.

Bottles of water, urine, and Lord knows what else.

I saw protesters armed with shovels, bats, hammers, and metal pipes.

Now, to be clear, I've marched in a few protests in my day.

But neither I, nor anyone that I was ever with, saw the need to bring weapons in order to lift up our voices and express our first amendment rights."

Um Count: 0

Sentiment Analysis: -5

A.2.4 Mayor 4

"Thank God there was no loss of life, there were no major injuries.

There was some real property damage, no doubt.

It was contained, and it was addressed consistently.

But, when I looked at the big picture here, situation that was very complex, very difficult, now unlike anything honestly we have seen in recent years.

And, and, this is something that bears real discussion and real analysis.

These protests were different, not because of the underlying issues, the underlying issues are profound and meaningful, again, expressed by those who are peacefully protesting and seeking change.

The acts factored here of a different kind of small set of now I would call them not just protests but people who came to do violence in a systematic, organized fashion."

Um Count: 0

Sentiment Analysis: -5

A.2.5 Mayor 5

"Why would you want to do anything to take away from this special, emotional moment for his family, and for his friends.

And so if you want to do the right thing, and if you're standing up and protesting for the right reasons, and when the family is saying publicly that violence and looting are not

consistent to what George would've wanted.

This is big George, gentle George as they have referred to him.

Why would anyone, whether you come from outside of the city or in the city, why would you want to do anything that would take away this moment, and the justice for George?"

Um Count: 0

Sentiment Analysis: -0

A.2.6 Mayor 6

"...of everything that we do, to be one people in every step of the life.

We're all flawed.

We know we have to make it better.

We work at it, but every member of our law enforcement, every member of each community: black, yellow, tan, any color, any faith, any, any belief in anything.

We are a family of one people and we work every day because we believe in love and in a greater power."

Um Count: 0

Sentiment Analysis: -2

A.2.7 Mayor 7

"Damage to small businesses here in our city, as if COVID-19 hasn't...hasn't given them enough challenges.

To break the windows of, of, *uh* businesses *um* that have nothing to do with the situation *um* is just, is just morally wrong.

And, and I think that needs to be stated *uh* and I hope that tomorrow we can have a much better protest.

I will say that tomorrow's protest is organized by local groups *uh* black lives matter and NAACP leadership that we know in this city that we have worked with for years.

Tonight's protest was not organized by local leaders that we know *uh* and so police have been concerned about it all day.

uh I'm very hopeful for tomorrow's protest *uh* and I hope we don't see a repeat of what we've seen tonight."

Um Count: 8

Sentiment Analysis: -3

A.2.8 Mayor 8

"Where we have *uh* peaceful, non-violent demonstrations, unfortunately for a few folks devolved into something different.

uh and so...we are working to make sure that we prevent *uh* the loss of life.

And also prevent any additional criminal behavior that would lose *uh* property and prevent or *uh* or *uh* jeopardize the safety of our neighbors.

So to that effect, *uh* I have instituted a curfew that began at 11:30 tonight and will last until tomorrow at 6am.

uh tomorrow at night there will also be a curfew from 10pm to 6am *uh* that will be in effect *uh* until Monday morning at 6am and for the central business district."

Um Count: 10

Sentiment Analysis: -3

APPENDIX B

DEMOGRAPHIC INFORMATION

B.1 Experiment I Data

		AMT	Snowball
Party	Rep	33	15
	Dem	41	45
	Ind	19	20
	Other	1	6
	NP	0	5
Age	18-29	19	46
	30-39	33	14
	40-59	32	13
	60+	8	20
Education	Less than high school degree	0	0
	High school graduate /GED	0	3
	Some college but no degree	13	13
	Associate degree in college (2-year)	8	3
	Bachelor's degree in college (4-year)	54	40
	Master's degree	18	26
	Doctoral degree	0	4
	Professional degree (JD, MD)	1	2
Race	White	74	83
	Black or African American	11	2
	American Indian or Alaska Native	0	0
	Asian	6	4
	Native Hawaiian or Pacific Islander	0	0
	Other	1	0
Gender	Female	33	45
	Male	60	44
	Gender-fluid/Non-Conforming	1	2
	Not Listed	0	0
	Prefer not to answer	0	0
Sexuality	Heterosexual	66	61
	Homosexual	9	11
	Bisexual	16	11
	Other	3	6
	Prefer not to say	0	0
Finances	Less than $20k	6	4
	$20,000-49,999$	31	28
	$50,000-74,999$	28	19
	$75,000-99,999$	18	19
	$100k or more	10	18
	Prefer not to say	1	3
Employment	Working (paid employee)	72	63
	Working (self employed)	14	10
	NW (temp laid off)	1	2
	NW (looking)	5	1
	NW (retired)	0	9
	NW (disabled)	1	0
	NW (other)	0	6
	Prefer not to say	1	0

Table B.1. Demographic information from both AMT and Snowball samples in Experiment I (US Governors and COVID-19).

B.2 Experiment II Data

		AMT	Snowball
Party	Rep	24	11
	Dem	41	45
	Ind	19	20
	Other	1	6
	NP	0	5
Age	18-29	19	46
	30-39	33	14
	40-59	32	13
	60+	8	20
Education	Less than high school degree	0	0
	High school graduate /GED	0	3
	Some college but no degree	13	13
	Associate degree in college (2-year)	8	3
	Bachelor's degree in college (4-year)	54	40
	Master's degree	18	26
	Doctoral degree	0	4
	Professional degree (JD, MD)	1	2
Race	White	74	83
	Black or African American	11	2
	American Indian or Alaska Native	0	0
	Asian	6	4
	Native Hawaiian or Pacific Islander	0	0
	Other	1	0
Gender	Female	33	45
	Male	60	44
	Gender-fluid/Non-Conforming	1	2
	Not Listed	0	0
	Prefer not to answer	0	0
Sexuality	Heterosexual	66	61
	Homosexual	9	11
	Bisexual	16	11
	Other	3	6
	Prefer not to say	0	0
Finances	Less than $20k	6	4
	$20,000-49,999$	31	28
	$50,000-74,999$	28	19
	$75,000-99,999$	18	19
	$100k or more	10	18
	Prefer not to say	1	3
Employment	Working (paid employee)	72	63
	Working (self employed)	14	10
	NW (temp laid off)	1	2
	NW (looking)	5	1
	NW (retired)	0	9
	NW (disabled)	1	0
	NW (other)	0	6
	Prefer not to say	1	0

Table B.2. Demographic information from both AMT and Snowball samples in Experiment I (US Governors and COVID-19).

APPENDIX C

STUDY QUESTIONS

C.1 Questions from Pilot Study

Possible answers to questions are located within the parentheses.

1. Are you 18 years of age or older? (Yes or No)
2. Are you a current resident of the United States? (Yes or No)
3. What is your state of legal residence? (All 50 states)
4. *Repeat x6* How trustworthy does the speaker sound? (Very trustworthy, trustworthy, somewhat trustworthy, neither trustworthy nor untrustworthy, somewhat untrustworthy, untrustworthy, very untrustworthy)
5. *Repeat x6* How likely is it that this person will provide you with factual information about the COVID-19 pandemic? (Very likely, likely, somewhat likely, neither likely nor unlikely, somewhat unlikely, unlikely, very unlikely)
6. *Repeat x6* How likely is it that this person will provide you with factual information during a natural disaster? (Very likely, likely, somewhat likely, neither likely nor unlikely, somewhat unlikely, unlikely, very unlikely)
7. Did you vote in the last election? (Yes or No)
8. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else? (Rep, Dem, Ind, Other, NP)
9. Do you think of yourself as closer to the Republican or Democratic party? (Dem or Rep)

10. What political party are you registered with, if any? (Rep, Dem, Ind, Other, None)

11. What year were you born? (Text Box)

12. What is the highest level of school you have completed or the highest degree you have received? (Less than high school degree, High school graduate (high school diploma or equivalent including GED), Some college but no degree, Associate degree in college (2-year), Bachelor's degree in college (4-year), Master's degree, Doctoral degree, Professional degree (JD, MD))

13. Are you Spanish, Hispanic, or Latino or none of these? (Yes or No)

14. Choose one or more races that you consider yourself to be: (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Other)

15. To which gender identity do you most identify? (Female, Male, Non-conforming, Not Listed, Prefer not to Answer)

16. Which of the following best describes your sexual orientation? (Heterosexual, Homosexual, Bisexual, Other, Prefer not to say)

17. Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that includes your entire household income in (previous year) before taxes. (Less than $20k, $20,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 or more, Prefer not to say)

18. Which statement best describes your current employment status? (Working (paid employee), Working (self employed), NW (temp laid off), NW (looking), NW (retired), NW (disabled), NW (other), Prefer not to answer)

19. If you have any comments, questions, or concerns from participating in this survey, please place them in the text box below. (Text Box)

C.2 Questions from Experiment I

Possible answers to questions are located within the parentheses.

1. Are you 18 years of age or older? (Yes or No)
2. Are you a current resident of the United States? (Yes or No)
3. What is your state of legal residence? (All 50 states)
4. What is 1+3? (2, 4, 6)
5. *Repeat x4* How trustworthy does the speaker sound? (Very trustworthy, trustworthy, somewhat trustworthy, neither trustworthy nor untrustworthy, somewhat untrustworthy, untrustworthy, very untrustworthy)
6. *Repeat x4* How likely is it that this person will provide you with factual information about the COVID-19 pandemic? (Very likely, likely, somewhat likely, neither likely nor unlikely, somewhat unlikely, unlikely, very unlikely)
7. Please select "Somewhat Trustworthy" in the answers provided below. (Very trustworthy, trustworthy, somewhat trustworthy)
8. Did you vote in the last election? (Yes or No)
9. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else? (Rep, Dem, Ind, Other, NP)
10. Do you think of yourself as closer to the Republican or Democratic party? (Dem or Rep)
11. What political party are you registered with, if any? (Rep, Dem, Ind, Other, None)
12. What year were you born? (Text Box)
13. What is the highest level of school you have completed or the highest degree you have received? (Less than high school degree, High school graduate (high school diploma or equivalent including GED), Some college but no degree, Associate degree in college (2-year), Bachelor's degree in college (4-year), Master's degree, Doctoral degree, Professional degree (JD, MD))
14. Are you Spanish, Hispanic, or Latino or none of these? (Yes or No)

15. Choose one or more races that you consider yourself to be: (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Other)

16. To which gender identity do you most identify? (Female, Male, Non-conforming, Not Listed, Prefer not to Answer)

17. Which of the following best describes your sexual orientation? (Heterosexual, Homosexual, Bisexual, Other, Prefer not to say)

18. Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that includes your entire household income in (previous year) before taxes. (Less than $20k, $20,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 or more, Prefer not to say)

19. Which statement best describes your current employment status? (Working (paid employee), Working (self employed), NW (temp laid off), NW (looking), NW (retired), NW (disabled), NW (other), Prefer not to answer)

20. If you have any comments, questions, or concerns from participating in this survey, please place them in the text box below. (Text Box)

C.3 Questions from Experiment II

Possible answers to questions are located within the parentheses.

1. Are you 18 years of age or older? (Yes or No)

2. Are you a current resident of the United States? (Yes or No)

3. What is your state of legal residence? (All 50 states)

4. What is 1+3? (2, 4, 6)

5. *Repeat x4* How trustworthy does the speaker sound? (Very trustworthy, trustworthy, somewhat trustworthy, neither trustworthy nor untrustworthy, somewhat untrustworthy, untrustworthy, very untrustworthy)

6. *Repeat x4* How likely is it that this person will provide you with factual information regarding the protests on police brutality? (Very likely, likely, somewhat likely, neither likely nor unlikely, somewhat unlikely, unlikely, very unlikely)

7. Please select "Somewhat Trustworthy" in the answers provided below. (Very trustworthy, trustworthy, somewhat trustworthy)

8. Did you vote in the last election? (Yes or No)

9. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else? (Rep, Dem, Ind, Other, NP)

10. Do you think of yourself as closer to the Republican or Democratic party? (Dem or Rep)

11. What political party are you registered with, if any? (Rep, Dem, Ind, Other, None)

12. What year were you born? (Text Box)

13. What is the highest level of school you have completed or the highest degree you have received? (Less than high school degree, High school graduate (high school diploma or equivalent including GED), Some college but no degree, Associate degree in college (2-year), Bachelor's degree in college (4-year), Master's degree, Doctoral degree, Professional degree (JD, MD))

14. Are you Spanish, Hispanic, or Latino or none of these? (Yes or No)

15. Choose one or more races that you consider yourself to be: (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Other)

16. To which gender identity do you most identify? (Female, Male, Non-conforming, Not Listed, Prefer not to Answer)

17. Which of the following best describes your sexual orientation? (Heterosexual, Homosexual, Bisexual, Other, Prefer not to say)

18. Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that includes your entire household income in (previous year) before taxes. (Less than $20k, $20,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 or more, Prefer not to say)

19. Which statement best describes your current employment status? (Working (paid employee), Working (self employed), NW (temp laid off), NW (looking), NW (retired), NW (disabled), NW (other), Prefer not to answer)

20. If you have any comments, questions, or concerns from participating in this survey, please place them in the text box below. (Text Box)

APPENDIX D

LISTS OF STIMULI SOURCES

D.1 List of Governors

State	State Num	Governor	Party	Born	Inauguration	Tiny URL Speech	Video Start Time
Alabama (AL)	1	Kay Ivey	R	10/15/1944	04/10/2017	tinyurl.com/y9v7f99g	0:40
Alaska (AK)	2	Mike Dunleavy	R	05/05/1961	12/03/2018	tinyurl.com/y9kl6tqh	1:23
Arizona (AZ)	3	Doug Ducey	R	04/09/1964	01/05/2015	tinyurl.com/yc9odo95	5:47
Arkansas (AR)	4	Asa Hutchinson	R	12/03/1950	01/13/2015	tinyurl.com/y7sxk2x7	8:06
California (CA)	5	Gavin Newsom	D	10/10/1967	01/07/2019	tinyurl.com/yahpy5jj	19:00
Colorado (CO)	6	Jared Polis	D	05/12/1975	01/08/2019	tinyurl.com/y9j4q7hb	13:37
Connecticut (CT)	7	Ned Lamont	D	01/03/1954	01/09/2019	tinyurl.com/ybcxmvmu	3:14
Delaware (DE)	8	John Carney	D	05/20/1956	01/17/2017	tinyurl.com/yd28bkhm	3:54
Florida (FL)	9	Ron DeSantis	R	09/14/1978	01/08/2019	tinyurl.com/ydc4smqn	0:05
Georgia (GA)	10	Brian Kemp	R	11/02/1963	01/14/2019	tinyurl.com/ydh2zfm9	18:49
Hawaii (HI)	11	David Ige	D	01/15/1957	12/01/2014	tinyurl.com/ya7yx48v	5:40
Idaho (ID)	12	Brad Little	R	02/15/1954	01/07/2019	tinyurl.com/y77onpgq	1:07
Illinois (IL)	13	J. B. Pritzker	D	01/19/1965	01/14/2019	tinyurl.com/y8gzyns2	12:48
Indiana (IN)	14	Eric Holcomb	R	05/02/1968	01/09/2017	tinyurl.com/ycjym4w6	17:38
Iowa (IA)	15	Kim Reynolds	R	08/04/1959	05/24/2017	tinyurl.com/yavf9qzj	3:18
Kansas (KS)	16	Laura Kelly	D	01/24/1950	01/14/2019	tinyurl.com/y98v8kqq	7:45
Kentucky (KY)	17	Andy Beshear	D	11/29/1977	12/10/2019	tinyurl.com/yb9trt6h	12:44
Louisiana (LA)	18	John Bel Edwards	D	09/16/1966	01/11/2016	tinyurl.com/yd5duwx3	9:10
Maine (ME)	19	Janet Mills	D	12/30/1947	01/02/2019	tinyurl.com/y72mk382	13:49
Maryland (MD)	20	Larry Hogan	R	05/25/1956	01/21/2015	tinyurl.com/ydxqsamt	28:50
Massachusetts (MA)	21	Charlie Baker	R	11/13/1956	01/08/2015	tinyurl.com/y9ycy6s7	3:32
Michigan (MI)	22	Gretchen Whitmer	D	08/23/1971	01/01/2019	tinyurl.com/yddkgg6c	9:12
Minnesota (MN)	23	Tim Walz	D	04/06/1964	01/07/2019	tinyurl.com/y8a4o4jr	8:53
Mississippi (MS)	24	Tate Reeves	R	06/05/1974	01/14/2020	tinyurl.com/y765z8kx	7:56
Missouri (MO)	25	Mike Parson	R	09/17/1955	06/01/2018	tinyurl.com/ycueess2	3:57

State	State Num	Governor	Party	Born	Inauguration	Tiny URL Speech	Video Start Time
Montana (MT)	26	Steve Bullock	D	04/11/1966	01/07/2013	tinyurl.com/y7cdxhpm	15:56
Nebraska (NE)	27	Pete Ricketts	R	04/19/1964	01/08/2015	tinyurl.com/y7zr79tv	2:26
Nevada (NV)	28	Steve Sisolak	D	12/26/1953	01/07/2019	tinyurl.com/yc926kuh	15:13
New Hampshire (NH)	29	Chris Sununu	R	11/05/1974	01/05/2017	tinyurl.com/y8shka58	9:30
New Jersey (NJ)	30	Phil Murphy	D	08/16/1957	01/16/2018	tinyurl.com/yc7yx327	13:!5
New Mexico (NM)	31	Michelle Lujan Grisham	D	10/24/1959	01/01/2019	tinyurl.com/y8k99e6u	37:23
New York (NY)	32	Andrew Cuomo	D	12/06/1957	01/01/2011	tinyurl.com/y8kcvgcq	7:04
North Carolina (NC)	33	Roy Cooper	D	06/13/1957	01/01/2017	tinyurl.com/yb5jr4l9	1:25
North Dakota (ND)	34	Doug Burgum	R	08/01/1956	12/15/2016	tinyurl.com/y93lh9ne	12:30
Ohio (OH)	35	Mike DeWine	R	01/05/1947	01/14/2019	tinyurl.com/ya234jbc	11:22
Oklahoma (OK)	36	Kevin Stitt	R	12/28/1972	01/14/2010	tinyurl.com/y9m6cjfd	0:01
Oregon (OR)	37	Kate Brown	D	06/21/1960	02/18/2015	tinyurl.com/yb7qaopp	3:30
Pennsylvania (PA)	38	Tom Wolf	D	11/17/1948	01/20/2015	tinyurl.com/y8ka5k2v	0:08
Rhode Island (RH)	39	Gina Raimondo	D	05/17/1971	01/06/2015	tinyurl.com/ydgormpy	20:01
South Carolina (SC)	40	Henry McMaster	R	05/27/1947	01/24/2017	tinyurl.com/y8w4hpfr	2:37
South Dakota (SD)	41	Kristi Noem	R	11/30/1971	01/05/2019	tinyurl.com/y7yyzj7b	3:07
Tennessee (TN)	42	Bill Lee	R	10/09/1959	01/19/2019	tinyurl.com/yb4rkzzu	6:13
Texas (TX)	43	Greg Abbott	R	11/13/1957	01/20/2015	tinyurl.com/y99ahfpb	19:50
Utah (UT)	44	Gary Herbert	R	05/07/1947	08/11/2009	tinyurl.com/y86tkuq9	7:54
Vermont (VT)	45	Phil Scott	R	08/04/1958	01/05/2017	tinyurl.com/ycaj56ko	3:05
Virginia (VA)	46	Ralph Northam	D	09/13/1959	01/13/2018	tinyurl.com/yajzkyk8	8:10
Washington (WA)	47	Jay Inslee	D	02/09/1951	01/16/2013	tinyurl.com/ycdygjzh	1:08
West Virginia (WV)	48	Jim Justice	R	04/27/1951	01/16/2017	tinyurl.com/y8fq62sb	11:19
Wisconsin (WI)	49	Tony Evers	D	11/05/1951	01/07/2019	tinyurl.com/y9ev6tgn	0:01
Wyoming (WY)	50	Mark Gordon	R	03/04/1957	01/07/2019	tinyurl.com/ybyaq7av	2:08

D.2 List of Mayors

Location	Mayor	Number	Party	Born	Inauguration	Tiny URL Link	Video Start Time
Los Angeles, CA	Eric Garcetti	1	D	2/4/1971	7/1/2013	https://tinyurl.com/yalrlzxv	1:47
Phoenix, AZ	Kate Gallego	2	D	10/21/1981	3/21/2019	https://tinyurl.com/yb24sqh8	1:45
Chicago, IL	Lori Lightfoot	3	D	8/4/1962	5/20/2019	https://tinyurl.com/yaoe6hl4	0:10
New York, NY	Bill de Blasio	4	D	5/8/1961	1/1/2014	https://tinyurl.com/yca2uynz	∼47:00
Houston, TX	Sylvester Turner	5	D	9/27/1954	1/2/2016	https://tinyurl.com/y6wxy39d	40:00
Las Vegas, NV	Carolyn Goodman	6	Nonpartisan	3/25/1939	6/6/2011	https://tinyurl.com/yd6bekbl	1:00
Oklahoma City, OK	David Holt	7	R	3/10/1979	4/10/2018	https://tinyurl.com/yafyo5oh	2:35
San Antonio, TX	Ron Nirenberg	8	I	4/11/1977	6/21/2017	https://tinyurl.com/yb8dpgcb	∼3:00

D.3 List of Historical Figures

Name	Tiny URL Link	Video Start Time
Richard Nixon	https://tinyurl.com/yav8x2qj	0:16
Bill Clinton	https://tinyurl.com/y82j5jqh	0:10
John F. Kennedy	https://tinyurl.com/ya4hfeoc	1:19
Martin Luther King Jr.	https://tinyurl.com/lw5hj8x	11:03