



## OPEN ACCESS

EDITED BY  
Oliver Niebuhr,  
University of Southern  
Denmark, Denmark

REVIEWED BY  
Claire Pillot-Loiseau,  
UMR7018 Laboratoire de Phonétique  
et Phonologie (LPP), France  
Juan Manuel Sosa,  
Simon Fraser University, Canada

\*CORRESPONDENCE  
Sara Pearsell  
✉ pearsels@mcmaster.ca

SPECIALTY SECTION  
This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Communication

RECEIVED 31 March 2022  
ACCEPTED 15 December 2022  
PUBLISHED 06 February 2023

CITATION  
Pearsell S and Pape D (2023) The  
effects of different voice qualities on  
the perceived personality of a speaker.  
*Front. Commun.* 7:909427.  
doi: 10.3389/fcomm.2022.909427

COPYRIGHT  
© 2023 Pearsell and Pape. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# The effects of different voice qualities on the perceived personality of a speaker

Sara Pearsell\* and Daniel Pape

Department of Linguistics and Languages, McMaster University, Hamilton, ON, Canada

Although previous studies investigated various aspects of voice quality perception and personality attribution there are no studies, to our knowledge, which simultaneously examine and compare the perception of various voice qualities when produced by the same individual. This work investigates how laryngeal and supralaryngeal voice quality variations of a speaker affect *listeners'* perceived personality traits (and thus perceived charisma) of that same speaker. Six Canadian English speakers produced paragraphs varying the following voice qualities: *modal*, *creaky*, *breathy* (natural and artificial), (hyper-)nasalization, and *smiling* (natural and extreme). Listeners of a perception experiment were then tasked to rate 10 statements for each presented audio stimulus. Statements were selected corresponding to a sub-section of the Big 5 personality traits shown to be linked to charisma perception. Results show significantly more positive listener ratings (i.e., higher ratings compared to *modal*) with medium effects sizes for both *smiling* variants across all personality traits. In contrast, *creaky* was perceived significantly more negatively overall for all personality traits, with a medium effect size. *Nasal* and *breathy* still achieved statistically significant rating differences compared to the *modal* baseline. However, the overall effect pattern was more complex, and effect sizes were small or negligible. Additionally, we found consistent differences for some voice qualities when examining listener ratings comparing male vs. female *speakers*: for both *creaky* and *smiling* (but not for other voice qualities), female speakers were rated more negatively when producing *creaky* for some personality traits, whereas both *smiling* variants were consistently rated higher for females compared to males.

## KEYWORDS

variability in speech, perceived personality traits, voice quality variation, speech perception, laryngeal and supralaryngeal influences on cognition

## 1. Introduction

Recently there has been an increased interest in analyzing and understanding the effects of speakers' voice quality differences and how these *produced* differences may impact the *perception* of these speakers' personality traits. Areas of interest in this topic range from clinical techniques for best practices for a healthy vocal production (while avoiding vocal strain) to popular culture tips to sound more professional or speak more effectively. One specific area of interest is the role of voice quality in the perception of a speaker's personality traits, and, more generally, how these traits relate to the perception of speaker *charisma*.

## 1.1. Voice quality vs. vocal quality

Voice quality has been defined as “the quasi-permanent quality of a speaker’s voice” (Abercrombie, 1967) and “those characteristics which are present more or less all the time that a person is talking. It is a quasi-permanent quality running through all the sound that issues from his mouth” (Abercrombie, 1967). Following this definition, and in line with researchers like Laver (1980) and Esling et al. (2019), voice quality differences are based on the specific auditory coloring of an individual’s voice as a result of the variations of both *laryngeal* and *supralaryngeal* features which continuously occur throughout an individual’s speech production. Several significant factors play a role for the variation of different laryngeal voice qualities: subglottal pressure (the air pressure below the vocal folds), medial compression (the contraction of the lateral cricoarytenoid muscles causing adduction; how tightly the vocal folds are pressed together), adductive tension (how tightly the arytenoid cartilages are pressed together at the posterior end of vocal folds) and longitudinal tension [the tension or slack of vocalis, thyroid and cricoid muscles, as well as the cricothyroid muscles (Laver, 1980)]. Following Laver’s research, the most common phonation types, or laryngeal settings, are (i) *modal* or *normal voice*, the baseline (and non-pathological) voice setting, (ii) *breathy voice*, which has a high rate of air flow during production, (iii) *creaky voice* (also known as *vocal fry*, *laryngealization* or *glottalization*) characterized by very low frequencies which can be irregularly timed, (iv) *harsh voice*, a speech pattern with a normal fundamental frequency but aperiodicity or noise in the spectrum, and lastly (v) *tense or strained voice*, produced with a low rate of air-flow (often described as a “metallic” voice).

To conceptualize voice quality, it is helpful to think of each voice quality as a landmark on a continuum, with *breathy* on one end of that continuum (produced with a more open glottis), and *creaky* on the other end (produced with a constricted glottis). *Modal* voice is found between these two extremes. As a general notion of *modal* voice, this vocal quality has a more regular and periodic vibration pattern; there is no audible friction of the vocal folds, and the muscular tension is moderate. The vibrations are regular along of the vocal folds, often characterized as a “neutral mode of phonation” (Laver, 1980, p. 110). Medial compression, adductive tension, and airflow from the lungs are all moderate, and the longitudinal tension is low (vocal folds are shorter and thicker). The described voice quality landmarks vary slightly between individuals but maintain the same directional proximity to one another (*breathy* on one end, *modal* more central, and *creaky* on the other end).

Although these laryngeal features are the most dominant aspect in the description of different voice qualities, both the Abercrombie (1967) and Laver (1980) frameworks include suprasegmental modification of non-laryngeal features such as retroflexion/retraction, smiling or nasality. In this paper, we adapt these definitions of voice quality, which are also

supported by the ANSI definition (i.e., that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar). Still, throughout this paper, we will use the term *vocal quality* to stress that this term would include both laryngeal features (e.g., modal, creaky, and breathy voice qualities) but also consistent and continuous suprasegmental feature variation (such as smiling and nasality, found on supralaryngeal and suprasegmental level). We hope that by using the term *vocal quality* we clearly define the inclusion of non-laryngeal vocal tract features since the term *voice quality* is often used very differently in the literature.<sup>1</sup> Please note that the suprasegmental features of interest in the current study, smiling and hypernasality, can of course co-occur with laryngeal voice quality variations. For the purposes of this study, the suprasegmental features of interest are produced with underlying modal voice quality, with modal voice representing the baseline measurement for each speaker.

## 1.2. Vocal quality and personality perception

Previous research has examined various acoustic features and perceptual cues and their relationship to personality trait attribution. Some of these studies have investigated the relationship between independent features of segmented speech signals such as  $f_0$  and pitch<sup>2</sup> (Puts et al., 2007; Rosenberg and Hirschberg, 2009; Quené et al., 2016; Berger et al., 2017), nasality and filled pauses (Möbius, 2003; Niebuhr and Fischer, 2019), amplitude or loudness (Novák-Tót et al., 2017), harmonics frequencies (Collins, 2000; Hodges-Simeon et al., 2010), and vocal quality (Wolf, 2015; Abdelli-Beruh et al., 2014) alongside their interaction with various personality traits. These results suggest that individual variation within physiological aspects of speech can play an important perceptual role in personality trait ascription.

In earlier research on the perception of vocal quality and perceived personality attributes, Pittam (1985) examined different vocal qualities of speakers and the impact of these qualities on listeners’ ratings of solidarity, attractiveness, and status of the speaker. This study found that listeners’ ratings of solidarity with a speaker were greater when there was the

<sup>1</sup> Our decision to use vocal quality instead of voice quality for this paper stems from discussions with other researchers who often defined voice quality as purely consisting of laryngeal differences.

<sup>2</sup> Pitch is a perceptual term taking into account the different acoustic properties of a complex acoustic waveform (normally consisting of the fundamental frequency and a number of optional harmonics). Therefore, pitch perception values can be different from measured fundamental frequency values.

presence of either *breathiness* or *whispery*<sup>3</sup> qualities in the speaker's voice. Perceptions of status were higher for *tense* voices as well as *breathy* voices compared to *whispery* and *nasal* voices (Pittam, 1985). In another study, Laver (1972) demonstrated an association of *breathy* voice with perceived higher sexuality and sensuality when the speaker was female but not when the speaker producing that *breathy* quality was male. Other studies have also demonstrated a correlation between certain vocal qualities and perceptions: the more significant the creakiness of a speaker, the higher the perceptions of that speaker's dominance or higher social status; the harsher the voice quality, the lower the perception of prestigious status (Esling, 1978; Scherer, 1979). Additionally, participants (who were described as young adults) rated voices with increased *creakiness*, above all the other vocal qualities assessed, as older. Esling (1978) and Scherer (1979) also suggest that this perception of age, as a result of the presence of *creakiness* in vocal quality, may account for the decrease in ratings associated with the friendliness and attractiveness of a speaker.

One major theory of personality and its associated traits is the *Big 5 of Personality Traits* (Norman, 1963; McCrae and John, 1992). Within this theory, personality traits are categorized and defined within five groups: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (or OCEAN for short). The framework for our personality trait perception is based on this theory resulting in a broader categorization of attributes such as attractiveness or speaker status into one or more of these five personality traits.

Despite the interest in vocal quality and personality perception, to our knowledge, the current research remains limited to the focus of between-subject designs. That is, different voice quality conditions were always confounded with different speakers, for example, examining *creaky* vocal quality and thus only the influences of that *creaky* voice quality on listener perception. Understanding both the perception and production of multiple, potentially influential vocal qualities an individual is consistently able to produce can provide insight into many areas of interest. These interests could range from (1) understanding how listeners perceive a multitude of possible variations within a speaker's productions, as well as (2) clinical opportunities for those suffering from pathologies impacting their productions, (3) to professional opportunities for those who are outside academia to improve the effectiveness of their speech productions and understanding how their voice and its productions are perceived by an audience. Furthermore,

<sup>3</sup> Whispery voice is categorized as a combination of glottal friction and voicing. This combination creates greater amounts of inter-harmonic noise, creating an almost flat spectrum with increased levels of energy (Laver, 1980). Breathy voice differs from whispery because of weaker medial compression and a decreased degree of voicing effort. However, Laver (1980) notes the perceptual boundary is not clear between whispery and breathy. In this paper we use the term breathy quality.

currently it remains unclear what aspects of vocal quality variation are most salient for the concept of a *charismatic* speaker (Signorello and Demolin, 2013).

The present study investigates vocal qualities varied in a within-subject design, focusing on the following vocal qualities: *modal*, *breathy*, *creaky*, representing opposing ends of the voice quality spectrum as well as a medial point between the two, and the additional qualities of *nasality* (specifically hypernasality), and *smiling*. Within this study, these vocal qualities are rated explicitly in terms of within-subject personality traits and more specifically in terms of charisma-related traits.

### 1.2.1. Creaky voice

*Creaky* voice, also referred to as *vocal fry*, *glottal fry*, *laryngealization*, and *glottalization*, has been extensively researched. *Creaky* voice can be categorized by its irregular vocal fold vibrations created by the amalgamation of high adductive tension, low longitudinal tension, high medial compression and low subglottal pressure (Laver, 1980; Ladefoged and Johnson, 2011). It usually occurs at the lower end of a speaker's f<sub>0</sub> range. Gick et al. (2013) explain that "In creaky voice, the vocal folds are very shortened and slackened to maximize their mass per unit length, and the IA (Inter Arytenoid) muscles are contracted to draw the arytenoid cartilages together. This action allows the vocal folds to stay together for a much longer part of the phonation cycle than in modal voicing..., only allowing a tiny burst of air to escape between long closure periods".

#### 1.2.1.1. Creaky voice and personality perception

Previous research remains equivocal as to the perceptual influence of *creaky* voice on a speaker's personality characteristics. One study by Yuasa (2010) found favorable listener impressions for increased usage of *vocal fry*, with associations to personality traits such as professionalism, genuineness, and nonaggressiveness, as well as other positive assumptions about a speaker (e.g., higher level of education). *Creaky* voice has also been associated with worthiness, intelligence and friendliness (Pittam, 1987). However, other studies contradict these results: Anderson et al. (2014) showed that the presence of *creaky* voice, specifically in women, has the potential to negatively impact ratings of education level, competence, and trustworthiness. Gobl and Ní Chasaide (2003) found that *creaky* voice represents impressions linked to boredom and sadness. *Creaky* voice is found to be dominant in both younger male and female populations (Wolk et al., 2012; Abdelli-Beruh et al., 2014). Despite being present in both genders, research has shown when it comes perception of creaky voice, female speakers are more frequently perceived negatively compared to male speakers (Anderson et al., 2014; Wiener and Chartrand, 2014; Pointer et al., 2022). Although these studies present conflicting results, personality traits selected across studies do not equate to the same meaning or interpretation. It

should also be noted, regardless of personality trait mismatching across studies, that gender (and perhaps context) appears to influence the perceptual impact of *creaky* voice on listeners, therefore providing insights for the hypothesized outcomes of the current study when varying speaker and/or listener gender.

### 1.2.2. Breathy voice

In voices which are considered healthy (i.e., non-pathological), *breathiness* is categorized by partial adduction along the length of the vocal folds, with both the medial compression and adductive tension at low values, thus resulting in the increased escape of air (Laver, 1980; Reetz and Jongman, 2020). The amount of air escaping during phonation can cause differences in the perceived *breathiness* of a speaker's voice, with less adduction and a more gradual closing of vocal folds making the voice sound breathier (Hanson, 1997).

#### 1.2.2.1. Breathy voice and personality perception

As previously described, *breathiness* has been shown to increase listeners' solidarity ratings<sup>4</sup> and perceived status (Pittam, 1985) as well as to influence perceived sexuality and sensuality for female speakers (Laver, 1980). However, research on the influence of this specific voice quality remains limited. Understanding the gap in the literature with respect to *breathy* vocal quality can provide further insight into how vocal qualities impact listeners' categorization of speakers' personality traits.

### 1.2.3. Nasal voice

*Nasality* is a vocal quality which results from nasal sound energy in the production of a speech signal. It is the result of the velopharyngeal port being either open or closed at inappropriate times or more than acceptable in a given language or dialect.

*Nasal* vocal quality is the acoustic result of the sustained and excessive coupling of the nasal and oral cavities during speech and can be categorized in one of two ways: hypernasality (i.e., going toward an excess of nasality) and hyponasality (i.e., going toward the absence of nasality). Hypernasality is caused by an excess of air leaking out through the nasal cavity when speaking. This results in extra (nasal) resonances in the acoustic speech stream. This type of nasality can be a result of several factors, from physiological issues, including structural problems (e.g., shortened soft palate or movement problem causing incomplete closure of the nasal cavity) to errors in sound acquisition (e.g., not learning, normally as a child, how to control the movement of air through the vocal tract cavities). Additionally, hypernasality still can have varying degrees of presence (more nasal and less nasal) and is primarily a result of both the size and status of the velopharyngeal port opening

<sup>4</sup> i.e., listeners' solidarity ratings with the perceived speaker, in other words the speaker currently being rated by the listener.

(Watterson and Emanuel, 1981; Warren et al., 1988); however, this is a separate factor from the presence or absence of nasality in speech production. Hyponasality is the opposite of hypernasality, in which not enough air can pass through the nasal cavity, resulting in a lack of nasal resonances in the speech signal as a result of a blockage or obstruction in the nasal cavity. This vocal quality is typical of the common cold (Tull, 1999).

#### 1.2.3.1. Nasal voice and personality perception

To our knowledge, there is no previous research on the perceptual impact of *nasality* variation (specifically hypernasality) in non-pathological voices, presenting a knowledge gap in the literature on this vocal quality and its effect on speaker perception. It is important to note that, in principle, hypernasality could be combined with other vocal qualities, such as *breathy* or *creaky* voice. In our study, we restrict our examinations to the effect of *nasality* coupled with underlying *modal* voice, thus excluding combinations with other phonation types. Furthermore, nasal coupling is *continuously* produced by means of a lowered velum throughout the full duration of a sentence/paragraph production.

### 1.2.4. Smiling

The physiological movements involved in *smiling* include the widening of the mouth, retraction of the lips, the lowering of the tongue dorsum, and the tendency of a speaker to lower their jaw (Shor, 1978; Erickson et al., 2009). As a result of these movements, the vocal tract shortens, therefore altering the auditory perception of a speaker through an increase in formant frequencies as well as amplitude (Tartter, 1980). Tartter found that *smiling* has an audible effect on speech, generally associated with increased positive interpretations in a *smiling* condition.

#### 1.2.4.1. Smiling and personality perception

A study by Vazire et al. (2009) explored the impact of the speaker's sex on the interpretation of listeners' *smiling* perceptions. The outcome of the study revealed two separate affective states, one for men and one for women. For women, *smiling* was viewed as a signal of trustworthiness and indicated warmth or enthusiasm to the listener. *Smiling* in men was interpreted as a lack of self-doubt, and increased confidence and calmness. Other research has found producing speech while smiling positively impacts speech perceptions, but has ceiling effects: excessive smiling does not increase the perception of charisma when compared to moderate smiling (Tschinse et al., 2022). For the present study, the inclusion of the *smiling* condition aims to reveal the connection between the effects of a *smiling* speaker on the perception of personality traits and effectiveness as a speaker when embedded in our experimental setup. Of particular interest for the current study, similar to the findings for *creaky* voice, is the mismatch in personality trait attribution when comparing (speaker) gender. Please note that

*smiling*, like hypernasality, could be combined with other voice qualities such as *breathy* or *creaky*. In our study, we will examine *smiling* only with an underlying *modal* voice.

These vocal qualities (modal, creaky, breathy, nasal, and smiling) have been examined individually and been ascribed personality trait correlates. As previously mentioned, there remains a lack of knowledge comparing these different vocal qualities, in combination, and across individual speakers. We hope to clarify the saliency of each of these vocal qualities when compared to each other, while simultaneously clarifying their interaction with respect to personality trait association.

### 1.3. Personality traits and charisma

The definition of charisma presented by Niebuhr and Fischer (2019) states: “*charisma is symbolic, emotional laden, and value-based communication style signaling leadership qualities such as commitment, confidence, and competence that affect followers’ beliefs and behaviors in terms of motivation, inspiration, and trust.*” To further understand how to conceptualize charisma and charismatic speech research has looked at listeners’ perceptual ratings for speakers’ voices. These ratings were obtained through a series of presented statements correlating to charisma which listeners would rate from positive to negative, depending on the statement of each scale (Rosenberg and Hirschberg, 2009; Tskhay et al., 2018). For example, Rosenberg and Hirschberg (2009) found that *charismatic speakers* were associated with the (personality) traits of being enthusiastic, charming, persuasive, and convincing, all traits which can be found and categorized within the *Big 5* (John and Srivastava, 1999). As there is increasing interest in the sources of perceived charisma and more generally influential speakers, relying on vague interpretations of *charisma* is insufficient while using only the *Big 5* of personality traits is too broad. By analyzing charisma within the traits of the *Big 5* a clearer and more concrete interpretation of charisma can be established. The motivation behind our research is two-fold. Firstly, quantifying charisma based on the *Big 5* allows for a targeted understanding of which attributes form different trait categories in order to create the concept of charisma, while concurrently allowing for a better “big picture” interpretation of personality traits perception using the *Big 5*.

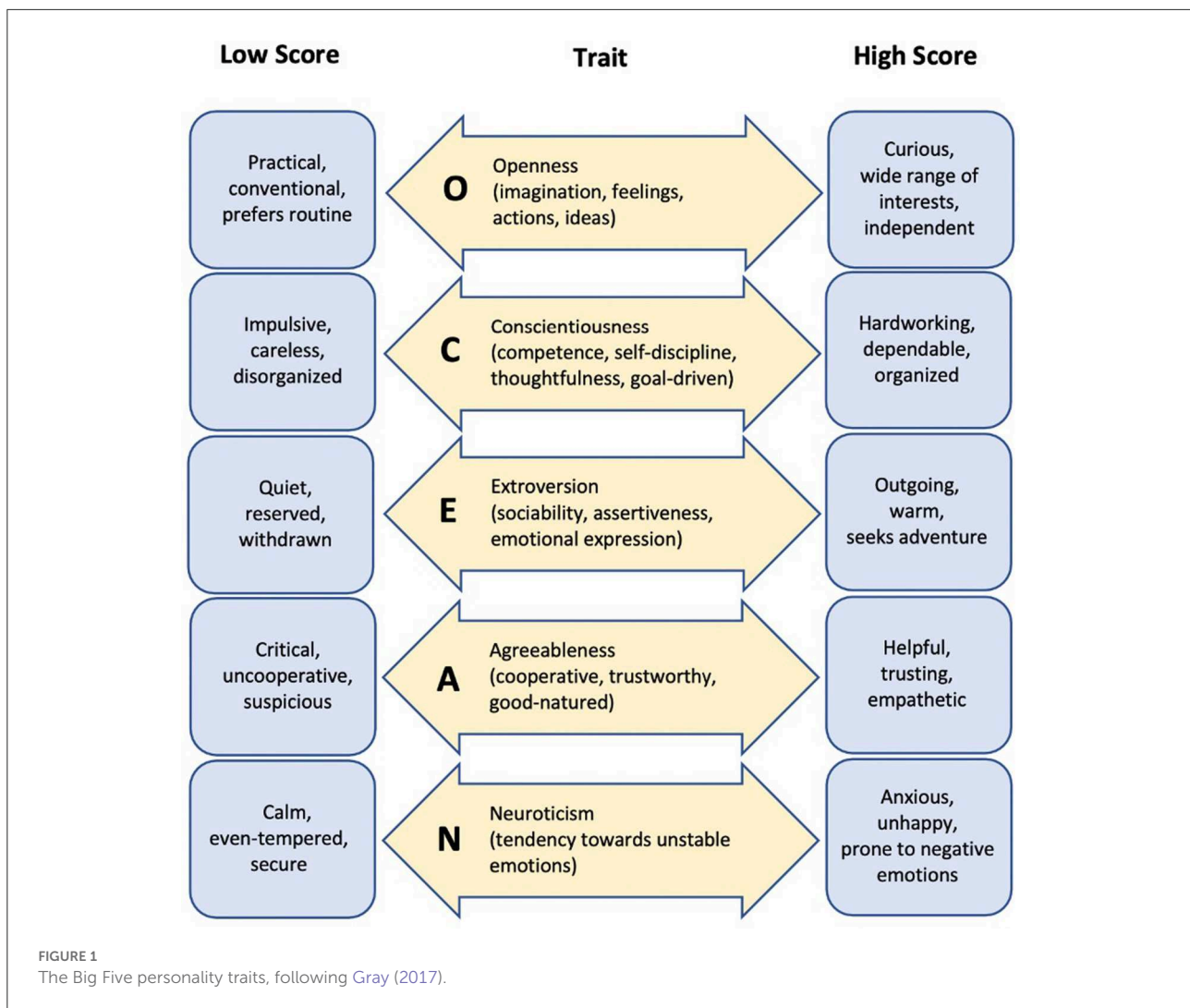
Although *charisma* may not be a trait in and of itself, there are still many personality traits that coincide with charismatic features of speech, as noted in a paper by Michalsky and Niebuhr (2019). As the authors point out, studies by several other researchers have demonstrated the relationship of the *Big 5* traits to charismatic speech features. Antonakis et al. (2016) implemented a training program targeted to teach charisma to managers and business leaders using a system called *Charismatic Leadership Tactics (CLTs)*. The purpose of these CLTs was to make the concept of charisma more tangible to learners. Within

their research, the authors demonstrate that confidence and self-assuredness are two facets which comprise charismatic speech. When examining these facets within the personality trait dimensions of the *Big 5*, these two facets fall into the *extroversion* personality trait (Costa and McCrae, 1992; John and Srivastava, 1999). Michalsky and Niebuhr (2019) also point out that the personality trait *agreeableness* relates to charismatic features, such as kindness, warmth, and development of trust while *conscientiousness* links to job performance and self-discipline (Costa and McCrae, 1992; John and Srivastava, 1999). Using just these examples, whether charisma is a personality trait in and of itself is debatable. Despite this, the traits associated with charismatic features of speech do have a relationship with personality traits and the *Big 5*, and exploration of charisma within the *Big 5* traits could provide a more general concept and understanding of the interaction of charisma perception and the use of vocal quality production.

In order to determine how different vocal qualities are attributed to the perceived personality traits of a speaker as well as how personality traits relate to charisma, the concept of personality traits needs to be further defined. As briefly mentioned above, one prominent theory of personality dimensions is that of the *Big 5 of Personality Traits* (Norman, 1963; McCrae and John, 1992). In this theory, personality traits can be described and categorized into the following sets: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism (or OCEAN for short). It is important to note that each of these categories is a range of extremes. For example, extroversion is on one side of the spectrum while introversion is on the other (John and Srivastava, 1999). There is a scoring system which takes participant responses to a number of questions and rates these responses as a score from high (e.g., extroversion) to low (e.g., introversion). Figure 1 provides a visual representation and brief summary of each of the five main traits as well as the traits associated with high and low scores.

The first trait is openness. This is a personality trait tied to imagination and insight, as well as openness to new experiences. Individuals with higher ratings in this trait are often perceived as more creative and have a wide-ranging set of interests. They are open, artistic, curious, and imaginative. Individuals who rate low in this trait are resistant to new ideas, are unimaginative, dislike change, and do not like to try new things (John and Srivastava, 1999). Using the questionnaires within the studies by Rosenberg and Hirschberg (2009), and Tskhay et al. (2018), we manually classified each of the questions presented there into the *Big 5 framework*. This later became the structure for our experimental design. From these two studies, the particular trait of openness has not been strongly associated as an indicator of charisma. In our experimental design, we therefore opted to omit this particular trait.

The second trait, conscientiousness, is linked to a person’s attention to detail, attentiveness, and goal-directed behavior. Those with a higher score in this trait are generally categorized



as efficient, organized, reliable, and responsible, while lower scores are associated with those who are less organized and more flexible in their approach to work. They may also procrastinate, lack discipline, and be careless, resulting in difficulty in completing tasks or goals (John and Srivastava, 1999). From the questionnaire list by Rosenberg and Hirschberg (2009), higher scores in conscientiousness related statements correlated to charisma but were however less proportionate in the number of statements presented than traits like extroversion and agreeableness.

Extroversion, the third trait, is related to the level and degree to which a person seeks interaction with their environment focusing on the social component. Those rating high in extroversion tend to be more social, assertive, outgoing, talkative, etc., while introversion, or those on the low rank of this trait, tend to be more reflective, and reserved, preferring solitude, avoiding being the center of attention and tend to be fatigued by an excess of

social interaction (John and Srivastava, 1999). Generally, a higher rating for extroversion is characteristic for charisma perception (Vergauwe et al., 2017), and extroversion is the Big 5 personality trait that receives the highest focus when determining charismatic attribution (Rosenberg and Hirschberg, 2009).

The fourth trait, agreeableness, determines how people treat their relationships with others. Unlike extroversion, agreeableness has to do with the pursuit of relationships with motivations concentrating on people's alignment and their interactions with others (John and Srivastava, 1999). Higher ratings in this trait indicate a person who is kind, forgiving, sympathetic, and trusting. Lower rating signal skepticism, stubbornness, a lack of sympathy, and a person who doesn't care about the feelings of others. The ability to connect with people as well as develop trust are just a few aspects which have also been demonstrative of charismatic speakers (John and Srivastava, 1999). Higher agreeableness scores appear to

also signal increased charisma in speakers (Rosenberg and Hirschberg, 2009).

The fifth and last trait is neuroticism. This is the trait which encompasses how an individual perceives the world, including the likelihood of inferring events as difficult or threatening as well as the inclination to experience negative emotions. People who rate high in this trait are anxious, tense, unstable, hostile, or irritable and experience dramatic shifts in mood. Those who rate lower are more emotionally stable, calmer, rarely feel sad or depressed, and do not often worry (John and Srivastava, 1999). In general, higher ratings of neuroticism have been shown to be negatively correlated with charisma and charismatic traits (Bono and Judge, 2004).

Although the Big 5 of Personality has traditionally been designed to be used by individuals based on introspection, the current study models questions/statements used in the previous studies of Rosenberg and Hirschberg (2009) and Tskhay et al. (2018). Within these studies, questions/statements were structured to be extrospective rather than the traditional introspective structure of Big 5 questionnaires. Other research (Hart and Hare, 1994; Ziegler et al., 2010) has demonstrated that ratings given by others fall closely within the range of ratings given from introspection. Theoretically, this means that results collected from our study's extrospective structure should produce data similar to those which would have been made by introspection.

#### 1.4. Aims of the study and hypotheses

As previously described, although there are studies which investigate individual vocal quality perception and personality attribution, there are no studies, to our knowledge, that simultaneously examine the perceptual effects of various vocal quality changes produced by the same individual speaker on the perception of charismatic traits within the context of the Big 5. The current research investigates how vocal quality variation (breathy, creaky, nasal, and smiling) of different speakers affects listeners' perceived personality traits and thus charisma of these same speakers. We are also interested whether one of these voice qualities is most salient in high (positive) vs. low (negative) personality trait ratings by listeners. Furthermore, we want to examine the influence of gender on listener perceptions: here we are interested in both the influence of *speaker* gender on listener ratings, but also the influence of *listener* gender.

Apart from the differences between vocal quality categories, we are also interested to examine the effects of two within-category modifications for *smiling* and for *breathy* voice. With respect to smiling, following the research by Tschinse et al. (2022) we are interested to either replicate or dispute the observed ceiling effect for *normal* vs. *extreme* smiling condition with our within-subject design, all with respect

to charisma ratings. With respect to breathy voice, we aim to introduce a technical, or more artificial, noise source modification in addition to the natural speaker-produced condition, thus examining the perceptual rating difference between a naturally produced breathy voice on the one hand vs. an artificially generated (technical) breathy voice on the other hand. The motivation here is to find out whether artificial noise added to the complete communication chain (and thus *not* modulated by laryngeal differences) would influence personality trait perception. In technical terms, the technical noise should be speech-shaped to make the conditions comparable and avoid adding another confound dimension.

We have the following hypotheses:

- H1: Lower listener perception scores, or negative ratings, for *creaky* voice across all speakers (resulting in a lower rating for all investigated personality traits, including neuroticism<sup>5</sup>), signaling a lack of perceived charisma in speakers. Lower scores for these traits in previous studies (Bono and Judge, 2004; Rosenberg and Hirschberg, 2009; Tskhay et al., 2018) have demonstrated a correlation to negative perceptions regarding speaker charisma.
- H2: Higher, or more positive, listener ratings in personality traits for *smiling*, with smiling having a positive correlation with speaker charisma.
- H3 (null hypothesis): Following Tschinse et al. (2022) we expect to see a ceiling effect for *smiling*, with the natural *smiling* productions expected to have almost identical rating scores compared to the extreme *smiling* condition.
- H4 (interactional hypothesis): We predict *speaker* gender to play a role in listener ratings. Specifically, female speakers will be rated more negatively when producing *creaky* voice (i.e., receive lower personality trait scores). For male speakers, we predict a less negative (or higher score) attribution of *creaky* voice compared to female speakers, thus bringing their ratings closer to *modal* voice ratings, meaning creaky voice for female speakers would be perceived less charismatically than their male counterparts.
- H5: For naturally produced vs. technical breathy voice we expect to see perceptual rating differences, with naturally produced breathy voice ratings lower for all examined personality traits and therefore rating lower in charisma. The reason for the lower expected ratings for natural breathy voice is that we assume that listeners are able to distinguish between noise as part of the speaker's laryngeal system (and thus being constantly modified by the speaker's production), whereas a channel-induced noise source could be better separated from the judged speaker characteristics,

<sup>5</sup> Please see section Experimental setup for the explanation about neuroticism scores.

and thus would influence personality perception ratings less than the natural breathy condition.

## 2. Materials and methods

### 2.1. Stimuli

We selected two paragraphs consisting of multiple simple sentences as the basis of the acoustic recordings to be used in the perception study. The paragraphs were constructed to have a neutral valence to prevent any impact from positive or negative valence in listener interpretations of the voice. Each paragraph was ~12 s long.

- 1) There is a house on the street and the kitchen door is open. Inside the kitchen, there's some table clothes in a basket. A spoon is on the table beside a coffee cup. I see a rug on the floor and magnets on the fridge.
- 2) The bedroom has two windows and a closet. A painting is hanging on the wall beside a clock. A dresser is across from the bed. Four drawers are in the dresser. There is a book and a lamp on the nightstand.

Six native Canadian English speakers (3 female, 3 male) recorded the paragraphs. Of these speakers, four were professional voice actors (2 male, 2 female), and two were Linguistics graduate students of McMaster University (1 female, 1 male). Due to lockdown restrictions associated with COVID-19, the four professional voice actors used their own high-quality microphones and adequate recording environments to record their productions and were directed and monitored *via* Zoom by the authors of this study. The two graduate student speakers recorded the stimuli using a high-quality microphone (Rode NT1A) and Focusrite Scarlet audio interface in the sound-proof booth of the Phonetics Lab at McMaster University. For all recordings, the microphone distance was specified to be around 10 cm, with the microphone being horizontally off-centre (from the lips) by ~30–45 degrees.

To ensure that the speech stimuli sounded natural without artificial manipulation or distortions, each of the voice qualities of interest was naturally (speaker-) produced. Although it can be challenging to produce several different vocal qualities on cue, we assumed that professional voice actors, as well as graduate students in Linguistics, would be highly skilled in their ability to do so. To ensure all speakers were producing exemplary productions for all vocal qualities and would sound highly natural, we explained each vocal quality, then acoustically demonstrated the vocal quality, and then continuously directed speakers on how to produce it. This included producing the vocal quality continuously throughout the produced sentences (i.e., from the start of the production to the end of the production), a comfortable and natural speech rate (not too

fast, or slow), as well as limiting pitch variation (as stable and flat  $f_0$  as possible), and amplitude variation (avoiding emphasis or stress). Once speakers were able to produce each vocal quality consistently and with the previously mentioned constraints (continuous vocal quality production throughout utterance, natural speech rate, stable and flat  $f_0$ , stable and consistent amplitude distribution) over the given paragraphs, they were then recorded. Both paragraphs were repeated three times for all voice qualities. The best of these repetitions was then selected as the stimulus for the listeners (i.e., the repetition with the least variation in pitch, and amplitude, continuous vocal quality production throughout the utterance, and natural speech rate). Prosodic differences were as tightly controlled as possible across speakers and conditions through the continuous direction during practice and recording sessions and auditory checks of the stimuli by the researchers. However, prosodic characteristics like  $f_0$  or intensity differences were not artificially manipulated to avoid the introduction of artifacts and did possess some variation across speakers. Since we are examining these vocal qualities against the speaker's own modal production (in other terms the baseline) we hope that any differences in prosodic control and variation across speakers' production is less impactful than if comparing directly to other speaker's productions.

The voice qualities produced were *modal*, continuous *nasalization* (specifically hypernasalization; hypernasality was produced with a lowered velum from beginning of a sentence to its end), continuous glottalization (*creaky* voice), continuous *breathy* voice, and continuous *smiling* classified into two conditions: natural (where speakers were instructed to produce a natural, comfortable smile while recording stimuli; labeled SmilingN), and extreme *smiling* (speakers were instructed to smile excessively and to an extreme while producing the stimuli; labeled SmilingEX).<sup>6</sup> The smiling conditions were also visually monitored during recording sessions. Within the *breathy* voice condition, we included two distinct classes: a *natural breathy* voice production (as produced by the speaker, labeled Breathyn) and an *artificial breathy* voice production (labeled BreathynT). This *artificial breathy* production was created by taking the measurement of HNR (Harmonic to Noise Ratio) of each speaker's *natural breathy* production and overlaying a speech-frequency shaped noise signal onto their *modal* production with identical HNR measurement as the *natural breathy* production but with a rather technical (or speech transmission channel) noise overlaid.<sup>7</sup> In total, the stimuli consisted of 7 different vocal qualities, including a *modal voice* production for each speaker as the baseline.

<sup>6</sup> Note that by "continuous" we mean produced from the start to the speech production to the end of the speech production.

<sup>7</sup> The HNR measurement was done using Praat's algorithm using the object type "Sound: To Harmonicity: (cc)".



After recording, the audio samples were screened with the audio editor Amadeus Pro (Hairer, 2021) and carefully checked for achieved accuracy and consistency of each vocal quality production by the two authors of this study. Additionally, a steep high pass filter (80 Hz for male speakers; 150 Hz for female speakers) was applied to remove and attenuate any additional low-frequency noise which may have been a part of the original recordings.

The final stimuli count was 84 acoustic stimuli (*six speakers x two paragraphs x seven voice quality conditions*). Stimuli were not repeated, so each acoustic stimulus was only played once for each set of questions.

In the following, we present results for measuring the acoustic parameters of the produced stimuli in the three vocal qualities *modal*, *creaky* and *breathy* to confirm that all stimuli were produced consistently and according to the specifications outlined above. The acoustic measurements used were average speech rate, average fundamental frequency, its standard deviation, CPP, HNR, jitter and shimmer. These measurements are presented in [Appendix A1](#). Generally, we found that both male and female speakers produce the stimuli in similar and expected ways. The average speech rate is approximately four syllables per second and does not vary systematically between the modal, creaky and breathy conditions. Furthermore, the average  $f_0$  and its standard deviation within speakers also remains consistent across conditions. Speakers show the expected decreases in creaky condition (Blomgren et al., 1998; exception: speaker EM) and very similar values for modal compared to breathy voice (except for speaker HK and to some extent MK). For breathy quality, decreased values for CPP (cepstral peak prominence) measurements are an indicator of breathiness in speech with smaller ratios representing greater differences in breathiness perception (Park et al., 2019; Murton et al., 2020). We avoided using HNR measurements as errors in location of individual pitch pulse onsets can strongly affect HNR (Hillenbrand, 1987). Since the parameter CPP strongly correlates with breathiness and is more resistant to errors in fundamental period location than HNR (Hillenbrand and Houde, 1996), we examined the CPP values and found all six speakers produce consistent differences between modal and breathy voice. With respect to creaky voice, the parameters jitter and shimmer are often used as acoustic correlates. Jitter and shimmer measure the acoustic irregularities of vocal fold vibration and are linked to roughness, hoarseness or breathiness of a voice with higher measurements correlating to increases to these aspects of speech (Blomgren et al., 1998). Specifically, jitter relates to frequency variation from cycle to cycle, while shimmer relates to amplitude variation (Murton et al., 2020). All of our speakers for almost all paragraph conditions demonstrate clear increases of jitter and shimmer in their creaky voice condition, corresponding to findings from Blomgren et al. (1998).

## 2.2. Participants

Twenty-seven participants took part in the perception study. They were primarily undergraduate students at McMaster University around the age of 20–23 (the majority of whom are studying Linguistics, Health Sciences or Psychology). All participants reported normal hearing and cognition. They answered a set of demographic questions including gender, age, acquired and spoken languages and musical education background. The experiment was conducted in a sound-proof booth at the Phonetics Lab at McMaster University using the Gorilla Experiment platform with wave file playback and using state-of-the-art acoustic playback conditions (Focusrite Scarlett audio interface, Sennheiser HD 598 linear frequency-response headphones). The duration of the experiment was around 60 minutes, including the pre-screening components.

## 2.3. Experimental setup

### 2.3.1. Scales

One effective way of eliciting the perception of personality traits is the use of continuous sliding scales. In voice quality and personality research, different researchers used very different types of scales. One study by Puts et al. (2007) examined the perception of dominance/authority through the use of scales to acquire ratings. The researchers posed questions to listeners about a speaker's voice, including the perception of a speaker's likelihood to win in a physical fight or the dominance or submissiveness of the speaker. Weiss and Moeller (2011) also utilized sliding scales to establish the likability of a speaker with the German antonyms *sympathisch—unsympathisch* (in English, a rough equivalent of *pleasant—unpleasant*). Several other studies (Rosenberg and Hirschberg, 2009; Berger et al., 2017; Niebuhr et al., 2018) have implemented statement-based questions, e.g., “The speaker is X,” with a study-specific decision of which perceptual qualities are selected for X. Among these studies, there are variations with these statement-based questions; some are simply yes/no responses, while for others, responses are presented as a sliding scale from *strongly agree—strongly disagree*.

### 2.3.2. Presented statements

The statements used in the present study were based on research by Rosenberg and Hirschberg (2009) and Tskhay et al. (2018). In their research, Rosenberg and Hirschberg selected tokens based on their own judgement on whether they perceived the token as being either *charismatic* or *non-charismatic*, resulting in 26 stimuli with a mean length of 10.09s. The tokens were as context neutral as possible (e.g., “It's a pleasure to meet with you today.”). For each of their 26 tokens, participants were asked directly to rate how charismatic the

sample was on a 5-point scale. Additionally, participants were then asked to rate additional 23 attributes, using statements based on previous literature on charisma (see below). Examining the Big 5 in relation to the questions presented by Rosenberg and Hirschberg and the research presented by Tskhay et al., openness was not a trait applicable to ratings of charisma and was therefore omitted in the current study. Due to the high number of vocal qualities in our study's design, 10 statements were presented rather than the original 26 of the Rosenberg and Hirschberg study to prevent an excessively long experiment, and these statements were first classified by personality trait type and then balanced according to the proportion of each personality type in Rosenberg and Hirschberg's study (five extroversion, three agreeableness, one conscientiousness, and one neuroticism). Modeling after previous research (Rosenberg and Hirschberg, 2009; Tskhay et al., 2018), the statements regarding the personality traits of agreeableness, extroversion and conscientiousness were designed to have higher scores of these personality traits corresponding to higher participant ratings, meaning ratings were more positively associated with that trait. For the neuroticism personality trait, the statement aimed to have lower scores for higher participant ratings. These lower scores have been positively correlated to charisma as high scores for this trait are often associated with more negative connotations such as anxiety and proneness to negative emotions; the higher the score for neuroticism, the more the trait is exhibited, the lower the score, the less the trait is exhibited. Despite all of our statements being framed positively [rather than both positively and negatively as in Rosenberg and Hirschberg (2009)], the results should not be skewed, as scores for the Big 5 relate to either high or low scores within each trait (openness, conscientiousness, extroversion, agreeableness, and neuroticism).

For the current study, these 10 statements were presented to the listeners. As described, these sentences were constructed by the researchers modeling the research of Rosenberg and Hirschberg (2009) and Tskhay et al. (2018) and created with a neutral valence to avoid any influence of positive or negative emotional connotations of the speech stimuli on listeners. The statements depicted the speakers as professors with the intention of establishing a relationship between the speaker and the (student) listener. Based off the previously mentioned results regarding the various vocal qualities, their uses, and the different speech environments or contexts in which they may be preferentially used, some of these vocal qualities within the current study might not be expected given the established context of an academic setting (i.e., that the speakers are "professors"). However, our rationale for labeling the speakers as professors was to prevent any other interpretations of social standing differences between different speakers, as well as between speakers and listeners, and speech environment (formal rather than colloquial like friends or family). All of these

conditions could impact the perceptual ratings of speakers. Since many of the participants were university students at McMaster University, we decided it would be both interesting and relevant to characterize speakers as professors.

The ten statements were split in time, with the first screen containing the first 5 statements and the following screen containing the last five statements (Figure B1 in Appendix). This was done to prevent participants from being overwhelmed by excessive text content on one computer screen. Each screen played the audio stimuli once. Participants would slide the "button" to the desired location on each scale (for each statement) to represent how much they either agreed or disagreed with each statement presented (Figure B1 in Appendix B).

The 10 statements were related to four of the Big 5 of personality traits:

- i) extroversion (5 statements like "This professor engages students in the classroom")
- ii) agreeableness (3 statements like "This professor is positive and likable")
- iii) conscientiousness (1 statement: "This professor is organized and detail oriented")
- iv) neuroticism (1 statement: "This professor is convincing in the way they speak").

The listener's task was to judge the ten presented statements with respect to the simultaneously and acoustically presented audio file (and thus containing the different recorded vocal qualities). As described before, each of the audio files contained one paragraph and was recorded by the 6 individual speakers with seven different voice qualities. The statements on the screen (which were accompanied by the presented audio stimulus) would then be judged by the listeners using continuous sliding scales (from *strongly agree* to *strongly disagree*). For our analyses, the *strongly disagree* end of the scale would be coded as 0% listener rating, and the *strongly agree* end of the scale would be coded as the 100% point of possible listener rating.

## 2.4. Statistical analysis

The statistical analysis was performed using the software R (R Studio Team, 2018) and RStudio (R Core Team, 2018). Parametric (e.g., ANOVA) or non-parametric (e.g., Wilcoxon) tests (depending on tests for normality of the data distributions) are used to determine whether listener responses/scores for each examined vocal quality (with respect to the presented statements and thus personality trait classification) as a dependent variable would be significantly different compared to the *modal* voice baseline (judging the exact same vocal quality stimulus and presented statement).

### 3. Results

#### 3.1. Speaker-specificity vs. vocal quality influences

First, we aimed to examine whether each examined speaker would indeed drive a *vocal quality difference* in participant ratings or whether participants instead chose to rate an overall and general speaker personality (i.e., a personality *gestalt*) independent of the presented vocal quality manipulations and variations. In other words, we wanted to examine if listeners indeed showed an influence of varying vocal qualities *for each speaker* or rather judged each speaker based on his/her overall *vocal personality*.

To examine this question, we first present each speaker, and each examined vocal quality for listener ratings using violin plots with overlaid boxplots, as shown in [Figure 2](#). Firstly, it can be seen that for each speaker, the median rating for *modal voice* differs, thus establishing an overall speaker effect on listener ratings. Additionally, certain speakers are judged more positively (i.e., achieve higher response ratings overall) than other speakers, e.g., speaker CS is rated more positively overall for all examined vocal qualities than, for example, speaker JF. Furthermore, the variation of vocal quality clearly shows an effect on listener ratings, with *creaky* voice obtaining consistently negative (lower) participant ratings (except for speaker EM) and *smiling* receiving consistently positive (higher) ratings (except for speaker SA). Please note, again, that for neuroticism, the scale is inverted as higher ratings for our scale correlate to negative neuroticism personality attributes. *Breathy* and *nasal* qualities show varying results compared to *modal voice* across different speakers, but it seems that their ratings rather closely correspond to the overall *modal voice* ratings for each speaker. Finally, there does not appear to be an influence from the speaker profession on listener ratings, as can be seen in [Figure 2](#): professional voice actors (EM, CS, SA, HK) do not show apparent rating differences compared to the Linguistics graduate students (JF, MK).

Next, we present in [Table 1](#) the correlation coefficients between *modal voice* and all other examined vocal qualities for each of the examined speakers and for all speakers combined. Theoretically, if listeners exclusively judge the acoustic personality of the underlying speaker (i.e., providing a rating of the speaker *gestalt* independent of the speaker-produced vocal quality), then correlations between modal voice and all other vocal qualities should be close to +1 (i.e., increasing *modal voice* ratings for that speaker should also increase all other vocal quality ratings simultaneously, thus excluding a possible effect of individual vocal quality on listener judgments). In contrast, if listeners exclusively judge the different vocal qualities (but choose to ignore the overall speaker identity), then correlations would strongly depend on the individual

vocal quality comparisons. For example, it could be expected, based on the results in [Figure 2](#), that the correlation between *modal* and *creaky* would be inversely related (i.e., closer to  $-1$ ) compared to the correlation between *modal* and *smiling* (which could be positively correlated closer to  $+1$ ), and all other comparisons showing varying correlations, but, most importantly, not being uniformly close to  $+1$  as this would suggest an absence of a judged vocal quality difference. The correlation coefficients in [Table 1](#) show that, for all 6 speakers, most vocal qualities obtain *varying* correlations (i.e., values *not* close to  $+1$ ), thus establishing a clear influence of vocal quality on all listener ratings. For example, speaker SA shows a very high negative correlation between *modal* and *creaky* voice (i.e., if ratings for this speaker's modal increase, the ratings for creaky decrease), whereas this speaker's correlation between *modal* and *natural smiling* vocal quality condition is highly positively correlated (increased modal ratings correspond to increased *natural smiling* ratings), which clearly shows the influencing effect of *creaky* compared to *natural smiling* vocal quality on listener ratings. However, when examining the correlation table, it can also be shown that the vocal quality correlations are rather complicated and not straightforward (e.g., correlations between *creaky* and *modal* are highly positive for five speakers, and *smiling* vs. *modal* is highly negative for two speakers), but, importantly, the table, together with [Figure 2](#), shows a clear influence of examined vocal qualities on overall listener ratings.

#### 3.2. Vocal quality influences

To examine statistical differences between the examined manipulated vocal qualities, we first conducted a Shapiro-Wilk test for normality of the data distributions for each of the vocal qualities (i.e., one test for all *modal voice* responses, one test for all *nasal voice* responses and so on). All tests for normality were highly significant (see density plots of the seven vocal qualities in [Figure C1](#) in [Appendix C](#); see also the distributions of each violin in the violin plot in [Figure 2](#)), so we cannot assume a normal distribution of the data and thus decided to conduct significance tests using the non-parametric Wilcoxon signed-rank (matched sample) test. We performed pairwise comparisons to determine the statistical significance (1) of each vocal quality compared to the baseline *modal voice* and furthermore (2) comparing the *natural breathy* vs. *artificially breathy* and *normal smiling* vs. *extreme smiling* vocal qualities. In sum, 8 Wilcoxon tests were conducted, and the significance values shown in [Table 2](#) are Bonferroni-corrected for these multiple comparisons. Effect sizes comparing each examined vocal quality compared to the modal voice perception are also reported. The table shows that *all* comparisons of the six examined different vocal qualities against the *modal voice* baseline are highly significant; thus, all 6 vocal qualities

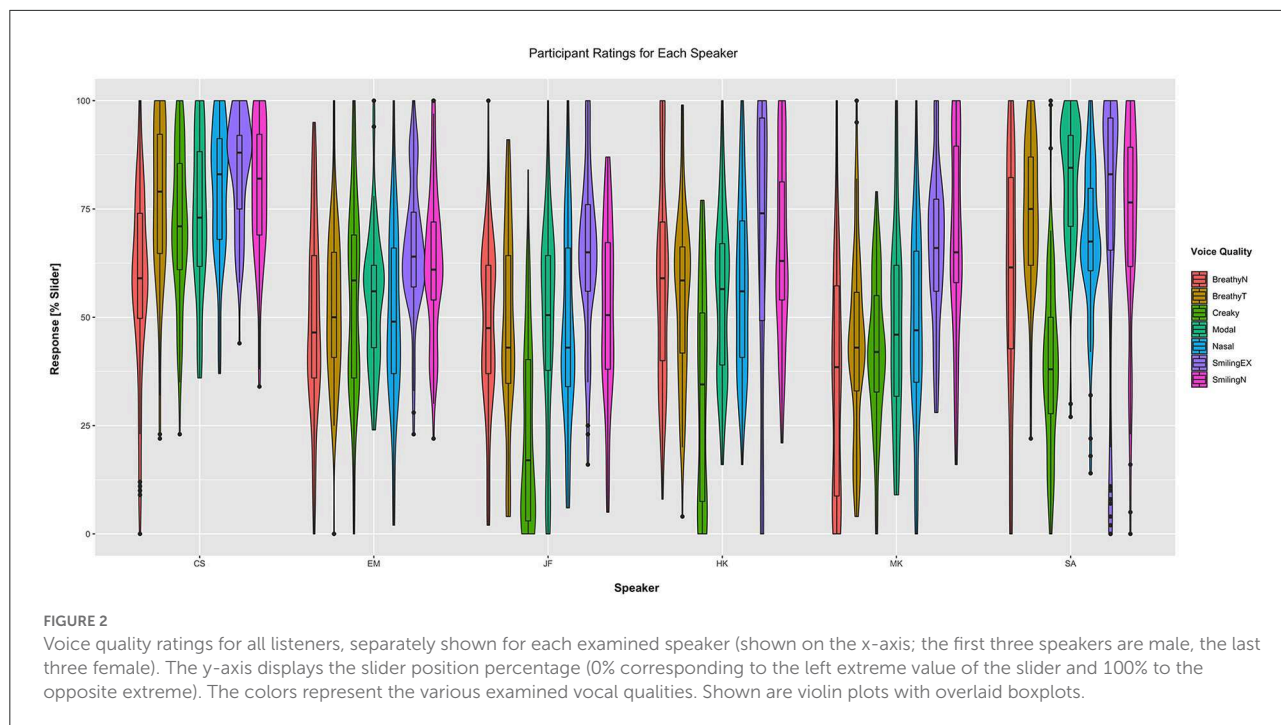


TABLE 1 Correlation coefficients for modal voice vs. all other examined vocal qualities (i.e., correlating Modal-Creaky, Modal-Nasal, Modal-BreathyNatural, Modal-BreathyTechnical, Modal-SmilingNatural, Modal-SmilingExtreme), calculated separately for each speaker and combined for the six speakers.

Speaker	Gender	Creaky vs. modal	Nasal vs. modal	BreathyN vs. modal	BreathyT vs. modal	SmilingN vs. modal	SmilingEx vs. modal
CS	Male	0.97	-0.21	-0.78	0.94	0.09	-0.79
EM	Male	0.85	0.80	0.70	0.99	-0.75	-0.65
JF	Male	1.00	0.85	0.53	0.93	0.79	0.73
MK	Female	0.72	0.32	0.29	0.96	0.26	0.40
HK	Female	0.68	1.00	0.97	0.97	1.00	0.68
SA	Female	-0.90	0.86	0.90	0.39	0.90	0.83
All speakers		0.47	0.42	0.00	0.91	0.28	0.02

obtain significantly different listener ratings when compared to the perceived modal voice baseline. Comparisons of the effect sizes show a medium effect size for both creaky (rated lower or more negatively compared to modal voice) and the two smiling conditions (rated higher or more positive compared to modal voice; with extreme smiling having a higher effect size). In contrast, the natural breathy condition has a small effect size (rated lower or more negatively compared to modal voice), and all other vocal qualities have negligible effect sizes. Finally, the pairwise comparison of the two smiling conditions shows for the Wilcoxon test that they are perceived significantly different, and the same is true for the comparison of the two breathy conditions, which also shows highly significant differences.

### 3.3. Personality traits vs. vocal quality influences

In the following, we aim to examine the interaction between vocal quality variation and perceived personality traits. Figure 3 shows violin plots with overlaid box plots over all speakers, split by the four examined personality traits (x-axis) and examined vocal quality (colors). The baseline would be the rating of the perceived modal voice, and it can be seen that this vocal quality shows very similar values when comparing the four personality traits. Visual examination of the vocal quality differences for each personality trait confirms the results of the previously presented significance tests: mean listener ratings were higher, or more positive, for smiling for both the natural smiling condition

TABLE 2 Mean, median and standard deviations for each vocal quality (the bold-printed vocal quality values of the first column are reported) and results of the statistical *Wilcoxon signed-rank pairwise significance test* (z-values: third column; p-values: fourth column) comparing the two vocal qualities stated in column one, calculated over all participant responses.

	Median, mean and standard deviation	z-values (Wilcoxon test)	p-values (Wilcoxon test)	Effect sizes (comparison to modal)
Modal	56, 53.6, 24.3	-	-	-
<b>Creaky</b> vs. modal	38, 38.8, 26	-26.99125	$p < 0.001^{***}$	-0.61
<b>Nasal</b> vs. modal	52, 50.9, 25.6	-6.021475	$p < 0.001^{***}$	-0.11
<b>BreathyN</b> vs. modal	50, 47.6, 26.1	-12.03835	$p < 0.001^{***}$	-0.25
<b>BreathyT</b> vs. modal	54, 51.6, 24.2	-4.951314	$p < 0.001^{***}$	-0.08
<b>SmilingN</b> vs. modal	67, 65.8, 21.6	-25.61255	$p < 0.001^{***}$	0.50
<b>SmilingEX</b> vs. modal	70, 68.2, 22.2	-28.47034	$p < 0.001^{***}$	0.60
<b>SmilingN</b> vs. smilingEx	-	-7.67	$p < 0.001^{***}$	-
<b>BreathyN</b> vs. breathyT	-	-8.37	$p < 0.001^{***}$	-

All p-values are Bonferroni-corrected. The last column gives the effect sizes (Cohen's d) for each examined vocal quality compared to modal voice perception. The \*\*\* symbol indicates statistically significant at the 0.1% level ( $p < 0.001$ ).

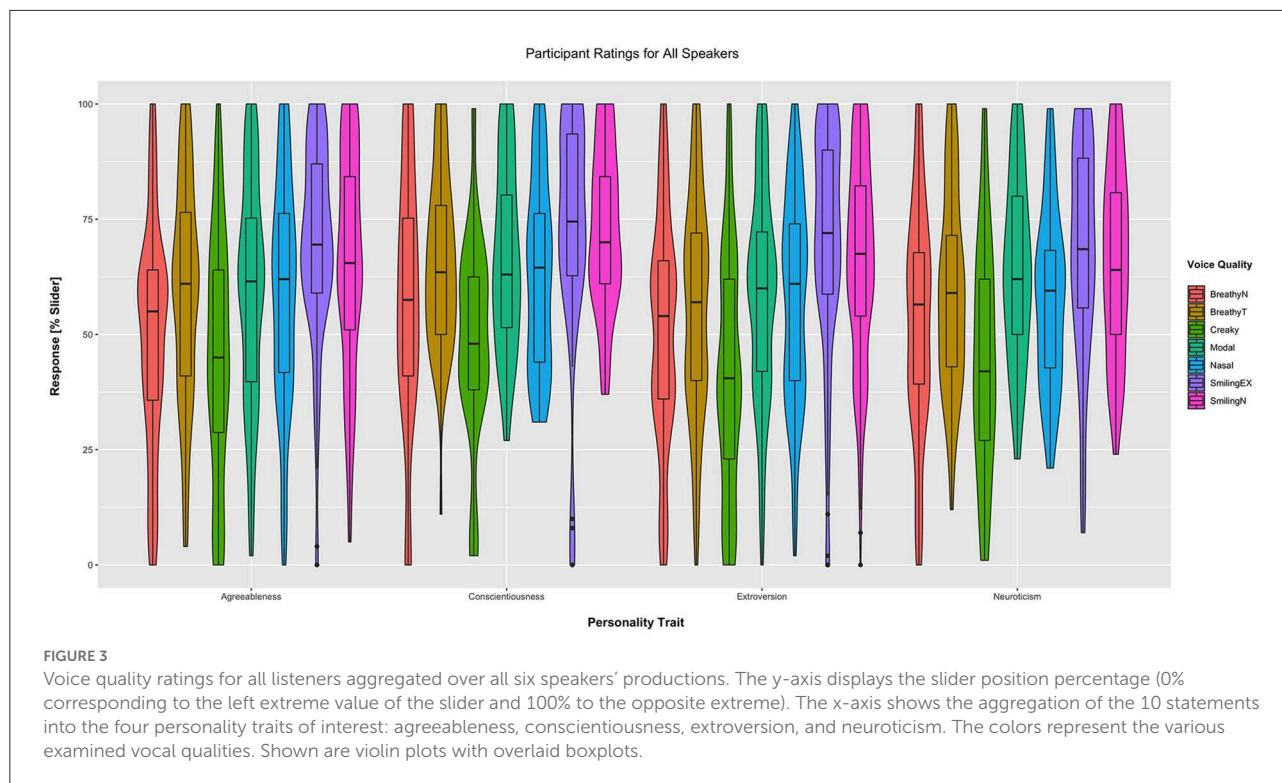
and the *extreme smiling* condition across all personality traits. Again, the *extreme smiling* condition is rated higher compared to the *natural smiling* condition, and the natural breathy condition is rated lower than the artificial breathy one, corresponding to the significant differences observed in Table 2. Inversely, *creaky* was perceived lower, or more negatively, for all four personality traits. This is in line with previous research (Tartter, 1980) that general perceptions of *smiling* are correlated to more positive emotions and associations like trustworthiness, friendliness, etc., while *creaky* is perceived more negatively.

While Figures 2, 3 show the differences between the seven examined vocal qualities and thus gave an appropriate first overview of the obtained listener responses and their response distributions, the main aim of this study is to investigate the difference between an observed vocal quality and its corresponding *modal voice* perception, or, in other words, to see the pure effect of each vocal quality manipulation with respect to the four personality traits. In order to see this effect, we calculated, for each vocal quality judgement, the *difference percentage* between examined vocal quality and the baseline *modal voice* for that exact same acoustic stimulus comparison, thus effectively providing a pure effect of each vocal quality on listener ratings, split by personality trait. For example, we took the judgement of listener 1 judging the first paragraph of speaker 1 produced in *modal voice* and subtracted this value from the judgement of listener 1 judging the first paragraph of speaker 1 in a *creaky voice*, thus providing a measurement value showing the absolute difference in vocal quality rating (compared to *modal voice* judgements) for that specific speaker, listener, and paragraph identity. This calculation was then performed for all other (vocal quality, speaker, and listener) judgements. Thus, this difference quantifies the effect of the magnitude of change in vocal quality

without taking into account other parameters. The results are presented in Figure 4, again as violin plots with overlaid boxplots. As can be seen, *smiling* again has the most considerable influence on all personality trait ratings, with more pronounced effects on agreeableness, conscientiousness and extroversion and a much smaller effect on neuroticism. In contrast, the *creaky* vocal quality has the strongest negative effect, with the most significant effect shown for conscientiousness compared to the other three traits. *Breathy voice* has a smaller negative effect on listener ratings, and interestingly this negative effect is strongest for neuroticism. An interesting result is the comparison of the two within-categories: the *natural smiling* vs. *extreme smiling*, and the *natural breathy* vs. *artificial breathy* condition. We do not observe the expected ceiling effect for *extreme smiling* conditions. Instead, for three of the four personality traits (excluding neuroticism), the *extreme smiling* condition consistently outperforms the *natural smiling* one, thus increasing the positive listener rating for more extreme smiling of each examined speaker. Interestingly, the *artificial breathy* condition generates more positive listener ratings compared to the *natural breathy* condition (see also Figures 2, 3 overall ratings), or, to turn it around, the *natural breathy* condition consistently leads to lower, or more negative, listener ratings compared to *artificial breathy* productions.

### 3.4. Effects of speaker and listener gender

When examining the effect of *speaker* gender, Figure 5 shows the mean differences in listener ratings, comparing listener judgements separately for male and female *speakers*, the produced vocal qualities and the four personality traits. We also provide results of the Wilcoxon pairwise significance



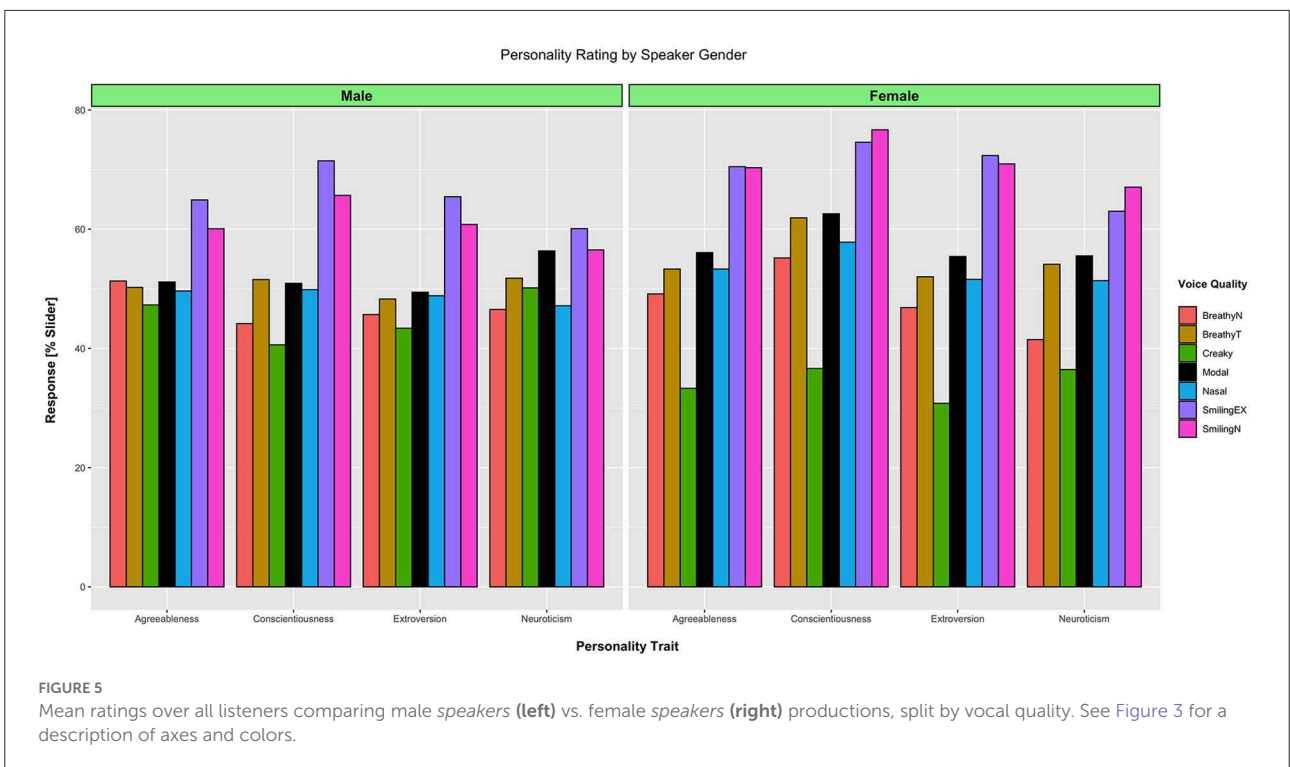
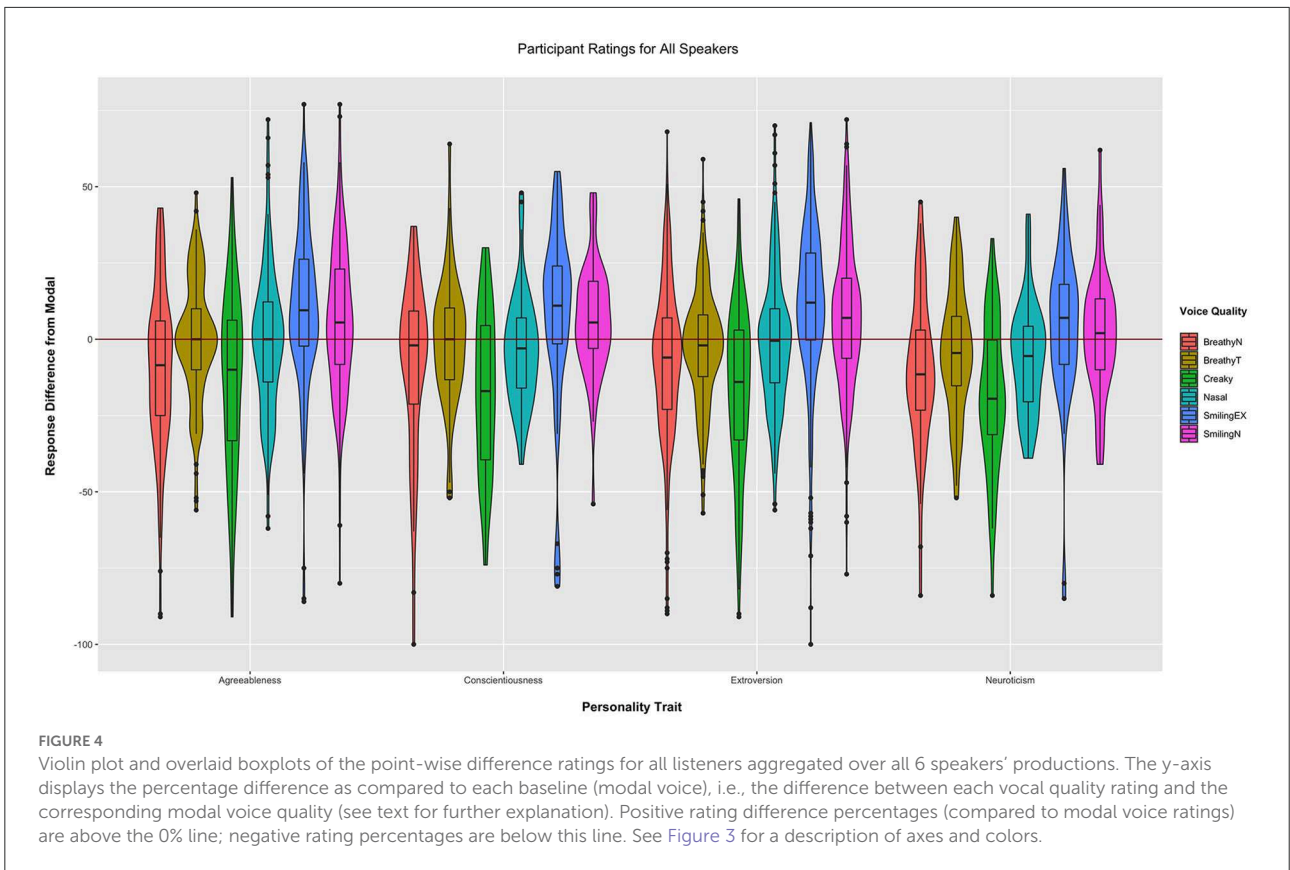
test in Table 3, over all speakers and for each examined vocal quality. Overall, the gender of the *speaker* has a significant effect on listener ratings ( $p < 0.0001$ ), and all vocal qualities except *natural breathy* voice show a highly significant effect of speaker gender (see Table 3). Examination of the means in Figure 5 shows that there is a tendency that female speakers are judged more positively, independent of the examined personality trait and for almost all vocal qualities. Furthermore, we consistently see larger differences for two vocal qualities: *creaky* and, to some extent, *smiling*. Listeners rated female speakers more negatively when producing *creaky* voice compared to males for all personality traits. Our results thus confirm previous research (Anderson et al., 2014; Chao and Bursten, 2021) that demonstrated that *creaky* voice is frequently perceived negatively in women in a variety of environments. Additionally, both *smiling* variants are consistently rated higher for female speakers than their male counterparts.

Figure 6 shows mean plots comparing the effects of male and female *listener* gender on personality ratings, split by vocal quality and personality trait. Table 3 provides the significance results for each vocal quality. Overall, *listener* gender, similar to *speaker* gender, also has a significant effect on listener ratings ( $p < 0.001$ ), however, the only vocal qualities rated significantly different when comparing the two (listener) genders are the *natural smiling* ( $p < 0.001$ ) and *artificial breathy* ( $p < 0.001$ ) vocal quality.

## 4. Discussion

Our results demonstrate that variation of the four different vocal qualities *breathy*, *creaky*, *nasal* and *smiling*, varied for the same individual speaker, can strongly and significantly influence listener perception of that speaker's personality traits, both positively and negatively: Results of the conducted Wilcoxon significance tests (see Table 2) show significantly higher listener rating scores for *smiling* voice qualities (for both *natural smiling* and *extreme smiling* condition) across all examined personality traits, whereas the *creaky* vocal quality was consistently and significantly rated lower for all personality traits for all participants, and thus perceived more negatively overall. For H1 we found that the continuous production of this creaky voice negatively impacts ratings of all personality traits. These results thus confirm the results of previous research regarding *creaky* voice and its unfavorable perception by listeners, which has linked this vocal quality to impressions of boredom and sadness (Gobl and Ní Chasaide, 2003), which would be classified into low neuroticism scores.<sup>8</sup> When specifically looking at *creaky* voice produced by women, the production of this vocal quality can negatively affect the ratings of competence (i.e., lower score in conscientiousness trait), trustworthiness (i.e., lower score in conscientiousness), and education level

<sup>8</sup> Based on our inverted scale for neuroticism described in the Methods section.



(Anderson et al., 2014). Furthermore, our results show this vocal quality is more negatively rated as compared to other vocal qualities such as *breathy* or *nasal* voice, both currently with very limited research results. Our results however are in contrast to those of both Yuasa (2010) and Pittam (1987), which found *creaky* voice correlated positively with professionalism, intelligence, friendliness, genuineness, and nonaggressiveness, as well as positive assumptions about a speaker, like assumed higher level of education. In sum, our study's lower ratings for all personality traits in combination with the results of previous studies (showing *decreased* scores for neuroticism, and conscientiousness being indicative of lack of charisma) all suggest that (continuously produced) creaky voice decreases the perception of speaker charisma. We can therefore accept this hypothesis.

For H2 our results are consistently and significantly higher ratings (see Table 2) for smiling. This confirms previous research, which found that *smiling* in women signals trustworthiness (high score in conscientiousness), indicated warmth (high score in extroversion trait) and enthusiasm (high score in extroversion trait) to the listener, whereas men who were *smiling* were interpreted as lacking self-doubt (high score in conscientiousness), confidence (high score in conscientiousness), and calmness (high score in neuroticism<sup>9</sup>; Vazire et al., 2009). This adds to general perceptions of *smiling* which are correlated to more positive emotions and associations like trustworthiness and friendliness (high score in extroversion; Tartter, 1980). Therefore, for H2 we can accept this hypothesis. Inversely to creaky voice, for *smiling* the observed higher ratings for all examined personality traits, in combination with the results of previous studies (showing *higher* ratings for neuroticism, extroversion, and conscientiousness being indicative of charisma), suggest that *smiling* positively impacts perceptions of speaker charisma.

For H3, also examining smiling, we find interesting results contrary to those presented by Tschinse et al. (2022). Our results in Table 2 show statistically significant differences between SmilingN (natural) and SmilingEX (extreme), with SmilingEX outperforming SmilingN with respect to positive listener ratings. These results thus reject our H3 null hypothesis. Although visually the differences between SmilingN (natural) and SmilingEX (extreme) appear rather small (see Figure 5), our statistical analysis demonstrates that increasing the smiling dimension also increases the positive influences on personality traits perception and therefore charisma. Whether the differences in results comparing our data with Tschinse et al. (2022) are due to our within-subject design or rather other methodological differences remains a cause for further study. Some of these mentioned methodological differences could be a result of stimuli: the stimuli in our study used short, isolated paragraphs (approximately 12 seconds) while the

<sup>9</sup> Again, based on our inverted scale for neuroticism.

TABLE 3 Results of the Wilcoxon signed-rank pairwise significance test (z-values and p-values) for each examined vocal quality and aggregated over all vocal qualities.

	Speaker gender		Listener gender	
	z-score	p-value	z-score	p-value
All (qualities)	-7.5408	<0.001***	-4.3274	<0.001***
Modal	-6.9697	<0.001***	-2.4725	0.01342
Creaky	-13.336	<0.001***	-2.1832	0.2367
Nasal	-4.976	<0.001***	-2.3980	0.01648
BreathyN	-1.1722	0.2391	-0.1885	0.8504
BreathyT	-5.4873	<0.001***	-3.74462	<0.001***
SmilingN	-14.0154	<0.001***	-3.40800	<0.001***
SmilingEx	-14.0154	<0.001***	-3.40800	0.2337

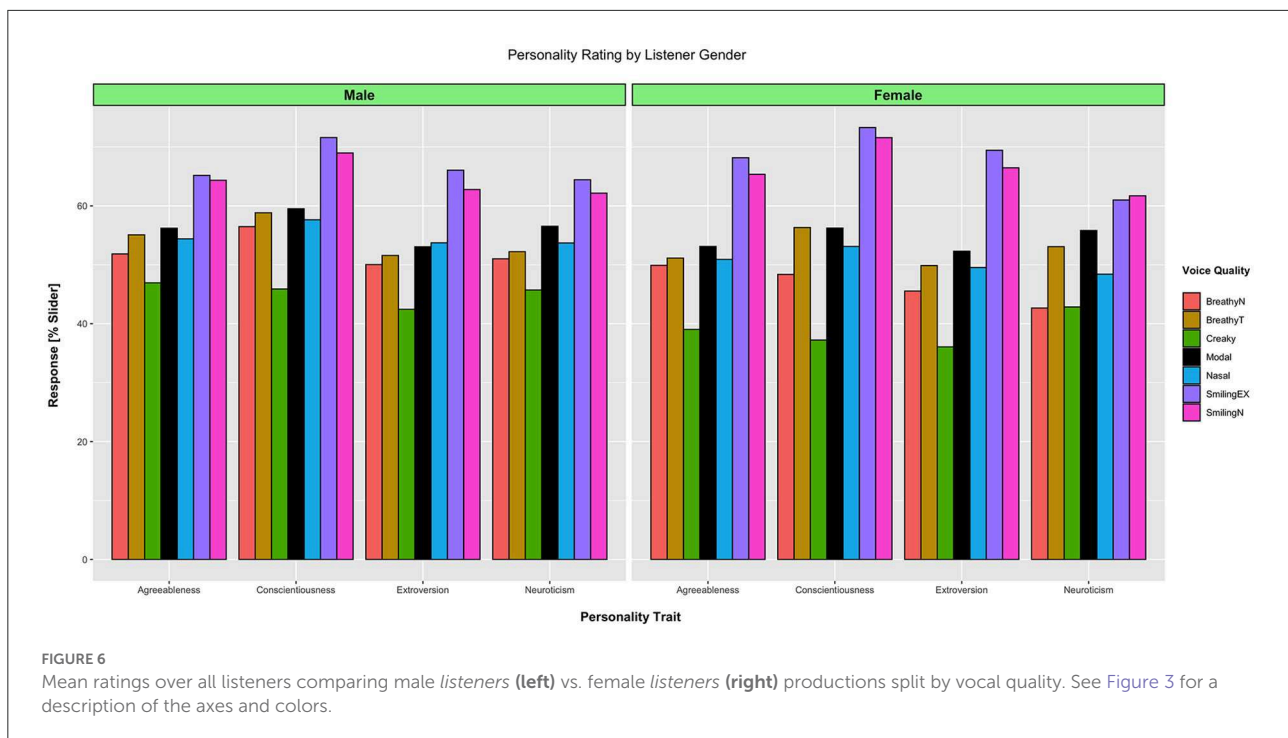
All p-values are Bonferroni-corrected (due to multiple comparisons). The \*\*\* symbol indicates statistically significant at the 0.1% level ( $p < 0.001$ ).

stimuli of Tschinse et al. (2022) were longer 1-min pitches. Prolonged auditory stimuli input allows for more habituation for participants and "saturation". Furthermore, the instruction for our extreme smiling and permanent smiling are not exclusively interchangeable, with the latter being temporally defined while the former is not.

For the other two vocal qualities, *nasal* and *breathy*, no consistent and robust differences in listener ratings across speakers could be found. However, the Wilcoxon significance tests showed that these two qualities still obtained statistically significant differences, all compared to the *modal voice* baseline (see Table 2). Despite this fact, our results for both *breathiness* and *nasality* do not suggest a strong and clear trend relating these voice qualities to individual perceived personality trait differences since overall ratings of *nasality* and *breathiness* follow very similar trends as the *modal voice* baseline. Our results thus suggest that both nasality and breathiness do not play a salient role compared to smiling or creaky voice for personality trait attribution, although they both seem to lower listener scores for several traits for most speakers. For *nasality*, this result is quite interesting as it fills a current gap in the literature regarding the saliency and influence of *nasality* on personality trait perception and charisma. For *breathy* voice, previous research has suggested that this vocal quality influences perceptions of a speaker's perceived sexuality and sensuality, but only when the speaker is female (Laver, 1980). Also, solidarity perception with speakers is higher for *breathy* voices (Pittam, 1985). From our data, the results suggest that *breathiness* does not have that strong influence (but again, please note the significant difference to modal voice based on the Wilcoxon test).

Together, the results show increases or decreases in listener ratings for each vocal quality type. They appear to either *all* increase or decrease together, depending on the positive or negative perception of that vocal quality. By





interpreting these traits collectively, we can see that those general increases/decreases of personality trait perceptions have a relationship with charisma; the higher, more positive, the ratings of traits, the greater the perception and saliency of charisma, whereas the lower, more negative the ratings, the lower the perception and saliency of charisma. Since we found that these increases/decreases in ratings synchronize across the different personality traits of the Big 5 (within each vocal quality), this can aid in future research on charisma in two ways: On the one hand, not all personality traits (of the Big 5) need to be utilized in experimental designs (i.e., only using questions/statements framed within the Extroversion trait, or Agreeableness, etc.) in order to capture meaningful interpretations of charisma and its presence. On the other hand, although statistically significant, some vocal qualities (nasal and breathy) are less salient in charismatic perception than other vocal qualities (creaky voice and smiling).

Further investigation into our within-category breathy voice differences reveal a statistically significant difference between BreathyN (natural breathy condition) and BreathyT (technical breathy condition). BreathyT is perceived with higher personality trait ratings and thus charisma, or, turning it around, BreathyN is perceived worse, thus confirming our H5. Despite being less salient than creaky vocal quality for charismatic trait attribution, these results suggest there is indeed a difference between adding the same type of noise (speech-shaped noise with identical HNR) to either the speaker's laryngeal signal (i.e., natural speaker-produced) or to the general communication channel (thus not being modulated by speech production

differences). We speculate here that listeners are indeed able to separate the added channel noise from the speaker (personality trait) judgments, thus pointing to a hypothesis that added channel noise is not as detrimental to personality trait perception as noise directly produced by the speaker's larynx.

With respect to perceptual saliency and the magnitude of participant rating differences, certain vocal qualities are more pronounced than others for personality attribution. For example, see the difference in point-by-point comparison of creaky vocal quality (compared to modal) vs. nasal vocal quality (also compared to modal) as shown in Figure 4. Although previous literature has shown the various impacts of how speakers are judged by listeners regarding vocal quality differences, the current results can provide a better understanding which vocal qualities may require more focus when attempting to increase charisma perception: smiling more and avoiding continuous creaky voice appear to be more relevant and more salient than avoiding nasal or breathy productions.

Furthermore, we examined the effect of speaker- and listener-dependent factors, namely the effect of gender for both speaker and listener. For H4, rating differences comparing male and female speakers and their vocal quality show consistent differences for both creaky and smiling vocal quality but not to the same extent for the other vocal qualities. Listeners rated female speakers more negatively in creaky voice than the corresponding male speakers for the personality traits of agreeableness, extroversion, and neuroticism. Here, our results confirm previous research (Anderson et al., 2014; Chao and Bursten, 2021) that demonstrated that creaky voice is frequently

perceived negatively in women in a variety of environments. Additionally, both *smiling* variants are consistently rated higher for female speakers than their male counterparts. In sum, our results confirm that gender strongly influences the perception of vocal quality, both overall and within different personality trait contexts and we can accept this hypothesis.

One limitation of the present study is the relatively small number of speakers and the observed inter-speaker variation (see, e.g., Figure 2, where listener ratings for speaker EM show the opposite pattern for *creaky* vs. *modal* voice compared to the other five speakers). Six speakers can provide a general picture of vocal quality and personality attribution, but this picture is still limited in the scope of potential variation, which may naturally occur in the production of individual speech patterns. Also, as can be seen in the correlation table, vocal qualities across speakers are not judged uniformly, thus introducing speaker-specific variation in this vocal quality study. For future studies, a higher number of different speakers could provide a more detailed understanding of the effect of vocal quality variations on charismatic traits, and it might continue to examine the more fine-tuned effects of *nasality* and *breathy* voice (in both technical and natural variation) that gave significant overall differences compared to *modal voice* perceptions but failed to provide a clear trend of the effects on individual personality traits.

An additional point of limitation in the current study is that we could not control for several other factors of variability for perceived vocal quality and personality ratings. For example, different possible settings would feed into the concept of *charismatic* speech and influence the ratings, for example: different communication contexts (formal vs. informal), environmental settings (e.g., academic, as in this study, vs. peer ratings), types of audiences (e.g., interviewers vs. colleagues), and of course whether the purpose of the communication is to be persuasive. Specifically, the established speaker-listener relationship in our study (i.e., the speaker being defined as a *professor* for our student participants) could influence vocal qualities to be perceived differently than if that relationship would have been established with a different social relationship paradigm (e.g., the listener is not a student, the listener is rating a friend, the listener is rating a co-worker, etc.). Since we chose this university setting—as previously explained—there is of course the chance that some of these voice qualities could be perceived differently for different speaker settings (friends or family) or different social environments (e.g., work, socializing with friends).

Finally, it is essential to note that other factors in combination with vocal quality appear to play a role, such as age vs. creakiness (Esling, 1978; Scherer, 1979) or the interaction of creakiness,  $f_0$ , and speech rate (Parker and Borrie, 2018). These, along with many other variations, suggest that, of course, vocal quality is not the only important component for listeners when giving ratings of personality traits. For future studies, the inclusion of other speech features like  $f_0$  variation, speech

rate differences etc., could provide a further understanding of the interactions between linguistic speech variation and voice quality.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving human participants were reviewed and approved by McMaster Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

SP and DP designed the experiment and stimuli, analyzed and interpreted data, and drafted and developed the manuscript. SP recorded the stimuli and collected participant data. DP also supervised the experiment. Both authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the NSERC Discovery Grant PGGPIN-2018-06518 to DP.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2022.909427/full#supplementary-material>

## References

- Abdelli-Beruh, N. B., Wolk, L., and Slavin, D. (2014). Prevalence of vocal fry in young adult male american english speakers. *J. Voice*, 28, 185–190. doi: 10.1016/j.jvoice.2013.08.011
- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Anderson, R., Klofstad, C., Mayew, W., and Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE*, 9, e97506. doi: 10.1371/journal.pone.0097506
- Antonakis, J., Bastardoz, N., and Jacquart, P. (2016). Charisma: an ill-defined and ill-measured gift. *Annu. Rev. Organ. Psychol. Organ. Behav.* 3, 293–319. doi: 10.1146/annurev-orgpsych-041015-062305
- Berger, S., Niebuhr, O., and Peters, B. (2017). “Winning over an audience - A perception-based analysis of prosodic features of charismatic speech,” in *Proceedings of the 43rd Annual Conference of The German Acoustical Society, Vol. 43 (DAGA)*, 1454–1457.
- Blomgren, M., Chen, Y., Ng, M. L., and Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *J. Acoust. Soc. Am.* 103, 2649–2658. doi: 10.1121/1.422785
- Bono, J. E., and Judge, T. A. (2004). Personality and transformational and transactional leadership: a meta-analysis. *J. Appl. Psychol.* 89, 901. doi: 10.1037/0021-9010.89.5.901
- Chao, M., and Bursten, J. R. (2021). Girl talk: understanding negative reactions to female vocal fry. *Hypatia*, 36, 42–59. doi: 10.1017/hyp.2020.55
- Collins, S. A. (2000). Men's voices and women's choices. *Anim. Behav.* 60, 773–780. doi: 10.1006/anbe.2000.1523
- Costa, P. T., and McCrae, R. R. (1992). *NEO PI-R Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Erickson, D., Menezes, C., and Sakakibara, K. I. (2009). “Are you laughing, smiling or crying?,” in *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference (Asia-Pacific Signal and Information Processing Association)*, 529–537.
- Esling, J., Moisis, S., Benner, A., and Crevier-Buchman, L. (2019). *Voice Quality The Laryngeal Articulator Model (No. 1)*. Cambridge, United Kingdom: Cambridge University Press. doi: 10.1017/9781108696555
- Esling, J. H. (1978). *Voice Quality in Edinburgh: A Sociolinguistic and Phonetic Study. Ph.D. Dissertation*. Edinburgh United Kingdom: The University of Edinburgh.
- Gick, B., Wilson, I., and Derrick, D. (2013). ‘Articulating Laryngeal Sounds’ in *Articulatory phonetics*. New York: John Wiley and Sons. p. 109.
- Gobl, C., and Ní Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* 40, 189–212. doi: 10.1016/S0167-6393(02)00082-1
- Gray, R. (2017). *Boundless Psychology*. Lumen. Available online at: <https://courses.lumenlearning.com/boundless-psychology/chapter/trait-perspectives-on-personality/> (accessed March 30, 2022).
- Hairer, M. (2021). *Amadeus Pro (version 2.8.8)*. London, United Kingdom: HairerSoft.
- Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *J. Acoust. Soc. Am.* 101, 466–481. doi: 10.1121/1.417991
- Hart, S. D., and Hare, R. D. (1994). Psychopathy and the Big 5: correlations between observers' ratings of normal and pathological personality. *J. Pers. Disord.* 8, 32. doi: 10.1521/pedi.1994.8.1.32
- Hillenbrand, J. (1987). A methodological study of perturbation and additive noise in synthetically generated voice signals. *J. Speech Lang. Hear. Res.* 30, 448–461. doi: 10.1044/jshr.3004.448
- Hillenbrand, J., and Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J. Speech Lang. Hear. Res.* 39, 311–321. doi: 10.1044/jshr.3902.311
- Hodges-Simeon, C. R., Gaulin, S. J., and Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nat.* 21, 406–427. doi: 10.1007/s12110-010-9101-5 nq
- John, O. P., and Srivastava, S. (1999). The Big Five trait taxonomy: history, measurement, and theoretical perspectives. *Handbook Personal.* 2, 102–138.
- Ladefoged, P. N., and Johnson, K. (2011). “Airstream mechanisms and phonation types,” in *A course in phonetics*. Boston etc.: Wadsworth.
- Laver, J. (1972). “Voice quality and indexical information readings,” in *Communication in face to face interaction*, Laver, J., and Hutcheson, S. (eds.). London, United Kingdom: Penguin. p. 189–203.
- Laver, J. (1980). The phonetic description of voice quality. *Cambridge Stud. Linguist. London*, 31, 1–186.
- McCrae, R. R., and John, O. P. (1992). An introduction to the five-factor model and its applications. *J. Pers.* 60, 175–215. doi: 10.1111/j.1467-6494.1992.tb00970.x
- Michalsky, J., and Niebuhr, O. (2019). Myth busted? *Challenging what we think we know about charismatic speech. AUC PHILOLOGICA* 2019, 27–56. doi: 10.14712/24646830.2019.17
- Möbius, B. (2003). “Gestalt Psychology meets phonetics-An early experimental study of intrinsic F0 and intensity,” in *Proceedings of the 15th International Congress of Phonetic Sciences, Vol. 3* (Barcelona: IPA and UAB), 2677–2680.
- Murton, O., Hillman, R., and Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *Am. J. Speech Lang. Pathol.* 29, 1596–1607. doi: 10.1044/2020\_AJSLP-20-00001
- Niebuhr, O., and Fischer, K. (2019). Do not hesitate!-unless you do it shortly or nasally: how the phonetics of filled pauses determine their subjective frequency and perceived speaker performance. *Interspeech*, 544–548. doi: 10.21437/Interspeech.2019-1194
- Niebuhr, O., Skarnitzl, R., and Tylečková, L. (2018). The acoustic fingerprint of a charismatic voice-Initial evidence from correlations between long-term spectral features and listener ratings. *Proc. Speech Prosody*, 9, 359–363. doi: 10.21437/SpeechProsody.2018-73
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *J. Abnorm. Psychol.* 66, 574. doi: 10.1037/h0040291
- Novák-Tót, E., Niebuhr, O., and Chen, A. (2017). A gender bias in the acoustic-melodic features of charismatic speech? *Interspeech*, 2248–2252. doi: 10.21437/Interspeech.2017-1349
- Park, Y., Perckell, J. S., Matthies, M. L., and Stepp, C. E. (2019). Categorization in the perception of breathy voice quality and its relation to voice production in healthy speakers. *J. Speech Lang. Hear. Res.* 62, 3655–3666. doi: 10.1044/2019\_JSLHR-S-19-0048
- Parker, M. A., and Borrie, S. A. (2018). Judgments of intelligence and likability of young adult female speakers of American English: the influence of vocal fry and the surrounding acoustic-prosodic context. *J. Voice* 32, 538–545. doi: 10.1016/j.jvoice.2017.08.002
- Pittam, J. (1985). *Voice quality: Its measurement and functional classification (Doctoral dissertation)*. Queensland University, Brisbane, QLD, Australia.
- Pittam, J. (1987). The long-term spectral measurement of voice quality as a social and personality marker: a review. *Lang. Speech*, 30, 1–12. doi: 10.1177/002383098703000101
- Pointer, N. F., van Mersbergen, M., and Nanjundeswaran, C. D. (2022). Listeners' attitudes towards young women with glottal fry. *J. Voice*. doi: 10.1016/j.jvoice.2022.09.007
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., and Gaulin, S. J. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evol. Hum. Behav.* 28, 340–344. doi: 10.1016/j.evolhumbehav.2007.05.002
- Quené, H., Boomsma, G., and van Erning, R. (2016). Attractiveness of male speakers: effects of voice pitch and of speech tempo. *Proc. Speech Prosody*, 8, 1086–1089. doi: 10.21437/SpeechProsody.2016-223
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/>
- R Studio Team (2018). *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc. Available online at: <http://www.rstudio.com/> (accessed January 15, 2022).
- Reetz, H., and Jongman, A. (2020). *Phonetics: Transcription, Production, Acoustics, and Perception*. New York: John Wiley and Sons.
- Rosenberg, A., and Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Commun.* 51, 640–655. doi: 10.1016/j.specom.2008.11.001
- Scherer, K. R. (1979). Voice and speech correlates of perceived social influence in simulated juries. *Lang. Social Psychol.* 5, 88–120.
- Shor, R. E. (1978). The production and judgment of smile magnitude. *J. Gen. Psychol.* 98, 79–96. doi: 10.1080/00221309.1978.9920859

- Signorello, R., and Demolin, D. (2013). The physiological use of the charismatic voice in Political speech. *Interspeech*. 987–991. doi: 10.21437/Interspeech.2013-174
- Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.* 27, 24–27. doi: 10.3758/BF03199901
- Tschinse, L. M., Asadi, A., Gutnyk, A., and Niebuhr, O. (2022). “Keep on smiling...? On the sex-specific connections between smiling duration and perceived speaker attributes in business pitches,” in *Proceedings 11th International Conference of Speech Prosody* (Lisbon), 1–5. doi: 10.21437/SpeechProsody.2022-127
- Tskhay, K. O., Zhu, R., Zou, C., and Rule, N. O. (2018). Charisma in everyday life: conceptualization and validation of the general charisma inventory. *J. Pers. Soc. Psychol.* 114(1), 131–152. doi: 10.1037/pspp0000159
- Tull, R. G. (1999). *Acoustic Analysis of Cold-Speech: Implications For Speaker Recognition Technology and the Common Cold*. Evanston, Illinois: Northwestern University.
- Vazire, S., Naumann, L. P., Rentfrow, P. J., and Gosling, S. D. (2009). Smiling reflects different emotions in men and women. *Behav. Brain Sci.* 32, 403–405. doi: 10.1017/S0140525X09991026
- Vergauwe, J., Wille, B., Hofmans, J., and De Fruyt, F. (2017). Development of a Five-Factor Model charisma compound and its relations to career outcomes. *J. Vocat. Behav.* 99, 24–39. doi: 10.1016/j.jvb.2016.12.005
- Warren, D. W., Hairfield, W. M., Seaton, D., Morr, K. E., and Smith, L. R. (1988). The relationship between nasal airway size and nasal-oral breathing. *Am. J. Orthod. Dentofacial.* 93, 289–293. doi: 10.1016/0889-5406(88)90158-8
- Watterson, T., and Emanuel, F. (1981). Observed effects of velopharyngeal orifice size on vowel identification and vowel nasality. *Cleft Palate J.* 18, 271–278.
- Weiss, B., and Moeller, S. (2011). “Perceptual dimensions of voice and way of speaking,” in *Electronic Speech Signal Processing (ESSV)*. p. 261–268.
- Wiener, H. J., and Chartrand, T. L. (2014). The effect of voice quality on ad efficacy. *Psychol. Market.* 31, 509–517. doi: 10.1002/mar.20712
- Wolf, N. (2015). *Young Women, Give Up the Vocal Fry and Reclaim Your Strong Female Voice*. The Guardian. p. 24.
- Wolk, L., Abdelli-Beruh, N. B., and Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *J. Voice.* 26, 111–116. doi: 10.1016/j.jvoice.2011.04.007
- Yuasa, I. P. (2010). Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women? *Am. Speech.* 85, 315–337. doi: 10.1215/00031283-2010-018
- Ziegler, M., Danay, E., Schölmerich, F., and Bühner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *Eur. J. Pers.* 24, 341–355. doi: 10.1002/per.753