# Influence of Vocal Cues on Perception of Traits: Evidence from Educational Context

**Mingyu Weng (weng97@purdue.edu)**
College of Education, Purdue University, IN, USA

**Wenxi Fei (wen-xi.fei@connect.polyu.hk)**
Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR

**Michael C.W. Yip (mcwyip@eduhk.hk)**
Department of Psychology, The Education University of Hong Kong, Hong Kong SAR

## Abstract

The human voice conveys more than just words; acoustic-vocal cues like pitch range and formant dispersion can shape perceptions of a speaker's personality. While research has explored this in various contexts, the impact of vocal cues on perceptions of teachers' traits remains unclear, particularly considering the educational level of listeners. This study investigates how college and secondary students perceive teacher utterances with manipulated acoustic parameters. Results showed that students from both age groups considered voices with a wider pitch range as being uttered by good teachers, but only the secondary students perceived a higher F0 and a wider formant dispersion as a feature of being a good teacher. Those suggest the mappings between teachers' characteristics and acoustic features might be different by age or education level, which could potentially the future teacher training for different levels of education.

**Keywords:** Education; Linguistics; Psychology; Attractiveness; Language understanding

## Introduction

Acoustic and vocal cues of human voice contain abundant information about the personal traits of speakers which can be perceived by listeners. Rather than being random, the relationship between the vocal cues and perceived personal traits is consistent. In fact, the inter-listener consistency in predicting the personal traits of speakers is significantly high across different experiment settings (McAleer & Belin, 2018).

In particular, pitch and intonation were found to play significant roles in the perception of trustworthiness. Pitch referred to the perceived fundamental frequency, while intonation was considered a pattern of the fundamental frequency. For example, in the studies on the relationship between pitch, intonation and the trait trustworthiness, listeners tended to feel the ascending intonation less trustworthy compared with the generally higher pitch levels (Tyler, 2015; Klofstad, Anderson, & Peters, 2012). It indicates that the perception of trustworthiness is associated with the pitch and pattern of intonation. Additionally, the relationship between formant dispersion and perceived dominance has been confirmed in previous studies. Formant dispersion refers to a function of mean distance among formants in human voices, which indicates the distance from the mouth to the larynx and the vocal tract length (Fitch, 1997). In former research, it was found that voices with smaller formant dispersion were perceived as being more dominant (Hughes, Harrison, & Gallup Jr, 2002; Puts, Hodges, Cárdenas, & Gaulin, 2007). Since the vocal tract length is a comparatively rigid parameter based on the physiological difference, it was argued that the traits of being dominant were judged based on the more innate biological features (McAleer, Todorov, & Belin, 2014).

When putting the interpretation of acoustic features in various social contexts, previous studies focused on the relationship between the perceived vocal traits and listeners' subsequent decisions and responses. Many studies have investigated how the perceived vocal cues affect judgment in situations, such as political elections (Klofstad, Anderson, & Nowicki, 2015), dating and mate selection (Xu, Lee, Wu, Liu, & Birkholz, 2013), and job interviews (Schroeder & Epley, 2015). For example, in the dating context, both male and female listeners tended to consider lower-pitch voices as more attractive (B. C. Jones, Feinberg, DeBruine, Little, & Vukovic, 2010; Xu et al., 2013), suggesting that listeners would take dominance, fertility, and physical body size as uttered by vocal cues into consideration in their mate selection (Xu et al., 2013; Borkowska & Pawlowski, 2011). In such social contexts, it is reasonable to argue that social factors such as gender and age may also influence listeners' perception of vocal traits. Previous studies suggest that the perceived voice traits can be related to the listeners' and speakers' gender. Scholars found that female listeners tended to feel the ascending intonation of female speakers was not trustworthy, but the same pattern by male speakers was perceived as more trustworthy (Tyler, 2015). Besides, the age of listeners also affects the perceived voice traits such as pleasantness and naturalness. Specifically, the older listeners were more likely to rate voices as pleasant, which differed subtly from the younger group who generally rated the same voices with more emotionally neutral scores (Goy, Pichora-Fuller, & van Lieshout, 2016).

In the context of classes or education in general, existing research reported mixed results about which and how acoustic cues play roles in making a speaker sound like a teacher with good personal traits such as reliability, trustworthiness, friendliness, and organizedness. Schmidt, Andrews, and McCutcheon (1998) found that the teachers with a narrower range of pitch levels in their voice seemed to be perceived with a more confident trait and believed by students to have the ability to improve the effectiveness of teaching. More recently, Gampel and Ferreira (2017) found that the teachers whose voices were equipped with a wider pitch range and



In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* 

loudness range were considered more pleasant by students. Thus, it suggests that students' perception of teachers' traits is likely to be affected by acoustic cues of the teachers' voices. Yet the relationship between acoustic cues and students' perception of teachers' traits requires more investigation since it seems that different acoustic cues were linked to different personal characteristics. Additionally, the age and educational level of different student groups may also affect their expectations of teachers when we consider the social factors (Constantinou & Wijnen-Meijer, 2022; Wei, Chow, Huang, Huang, & Cheng, 2023). Although many studies investigated how people from different age groups estimated the speakers' ages based on their vocal cues (Waller, Eriksson, & Sörqvist, 2015), Oosterhof and Todorov (2008) found that younger listeners, especially adolescents, tended to perceive a voice as trustworthy than older listeners, indicating the listeners' age can affect the perception of speakers' traits as well.

However, as we have reviewed, only limited studies focused on this particular domain of education together with students' perceptual perspectives and the age effect on perceiving teachers' traits. Thus, the present study wanted to investigate how vocal cues affect students' perception of whether a teacher is generally a good teacher and how these vocal perceptions differed between secondary and college students. We hypothesized that two age groups would show different preferential patterns when they were asked to judge the traits of teachers based on their vocal cues as suggested by previous studies on age effects. In sum, this study attempted to link acoustic features, vocal perception of traits, and age as a social factor together to explore how teachers can perform better or in a better vocal image to students in their classrooms.

# Methods

## Participants

Two age groups of participants, Group A and B, participated in this study. For Group A, 16 secondary school students (7 males) were recruited, aged between 12 and 15 ($M = 14.38, SD = 0.72$) at the time of the experiment. For Group B, 16 undergraduate students (7 males) were recruited, they were older than Group A (aged between 20 and 25, $M = 23.12, SD = 1.17$) and none of them were secondary school teachers. Participants from both groups were native speakers of Mandarin Chinese and learned English as a main subject from their primary school at about 7 years old. All of them received their primary and secondary education in mainland China. They also completed the Language History Questionnaires (Li, Zhang, Yu, & Zhao, 2020) to put the unexpected influence of their language background and language proficiency under control. Note that the mean length of studying English is 8.31 years ($SD = 0.87$) for Group A, and 17.29 years ($SD = 1.21$) for Group B, longer than Group A because they received a longer education at school. None of the participants in either group reported any history of serious brain illness or damage, hearing loss, language difficulty, or learning problems. Each participant was voluntary for the approximately 15-minute experiment. The participants were recruited on a voluntary basis and were acknowledged that they have the right to withdraw from the experiments at any time with no negative influence. No interest conflicts were reported. The study was approved by the Human Research Ethics Committee of the Education University of Hong Kong.

## Stimuli

Since the complexity of the intonation system of stimuli would potentially influence the perception (Scherer, 1972), English was chosen for the stimuli due to its relatively smaller phonological inventory compared to Mandarin Chinese (Lee & Zee, 2003; Roach, 2004). The vocal stimulus was synthesized using VocalTractLab (Birkholz, 2013). It was suggested that the listeners tended to make a judgment on personalities within a very short time (McAleer et al., 2014). In fact, listeners decided whether the speakers were nice or not from simply a brief utterance (McAleer & Belin, 2018). The relationship between short exposure and perception also explains the reason for the oversimplified perception of personality. In this study, the synthesized single word "bitter" with two syllables is chosen as the stimulus with an approximately 500ms duration, given our pilot experiment suggested that this utterance sounds less machine-like compared to others. Considering that it is a single utterance with no sentence context, the internal meaning of the word would have the least influence on the listener's perception. That duration was determined because a 300ms exposure to each stimulus was proved sufficient enough for people to make a judgment of vocal traits (McAleer et al., 2014).

Three acoustic parameters were manipulated in this study: fundamental frequency (F0), pitch range, and formant dispersion. For both F0 and pitch range, the stimuli were resynthesized, adopting the Manipulation function in Praat (Boersma, 2011). Three levels of F0 (90Hz: Low, 100Hz: Mid, 110Hz: High) and three levels of pitch range (50%: Low, 100%: Mid, 150% wide: High) were manipulated based on a pilot experiment that determined the minimal difference which could be perceived by the participants. For formant dispersion, there were three levels of formant dispersion, with the formant shift ratio being manipulated to 95% (Low), 100% (Mid), and 105% (High), using the *Convert* function in Praat. In total, there were 54 conditions of voices, including both male and female voices (3 conditions of F0 × 3 of pitch range × 3 of formant dispersion × 2 genders). Figure 1 is an example comparing stimuli of different parameters, from left to right: high (H) F0 (1) + medium (M) pitch range (2) + medium (M) formant dispersion (3) (H1-M2-M3), L1-M2-M3, and L1-H2-M3. To avoid the effects of acoustic steps (i.e. the comparison between high and low has a larger step than high and medium), we only considered the comparison pairs of the adjacent ones (for example, high (H) - medium (M) and medium (M) - low(L)). For each trial, there was only one factor manipulated within one acoustic step provided between two stimuli.

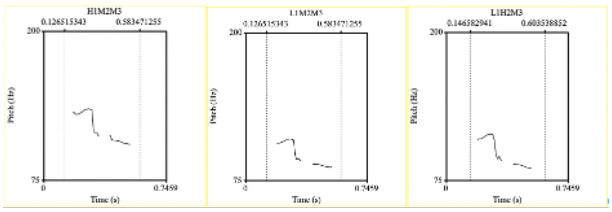


Figure 1: Exemplar stimuli with manipulated fundamental frequency and pitch range.

## Procedure

The experiment was built using PsychoPy (Peirce et al., 2019) and was conducted on a laptop (MacBook Air 2020) with a headphone (Bose QC35) in a quiet room. The two-alternative-forced-choice (2AFC) method was used in the experiment to ask the participants to choose one sound that fit the described trait better from two voices.

The participants were clearly given the context that the speaker to be judged is a secondary school teacher. In each trial, participants would see a fixation for 500ms first, followed by the targeted question about the expected trait to be evaluated for 500ms (e.g., "Which one sounds more like a responsible teacher?"). After that, participants were required to listen to two voices (around 500ms each) with a 300ms pause in between and choose the voice by pressing the F and J on the keyboard. Those target questions including those on trustworthiness, friendliness, and organizedness, were adapted from the Schmidt et al. (1998)'s list, and we also added a general question on *teachers' goodness* by asking "Which one sounds more like a good teacher?".

For each trial, only one targeted acoustic parameter out of the three will be modified in order to control the variables (e.g., low F0 - low pitch range - low formant dispersion: L1-L2-L3 vs low F0 - low pitch range - medium formant dispersion: L1-L2-M3). A training session was provided, including five trials, for the participants to understand the procedure. After the training session, the student would be given 162 trials to judge with 54 possible combinations of parameters. The stimuli were played in completely random order. During the experiment, there were six vigilance trials to examine the participants' attention. All participants completed the vigilance trials with all answers correct.

## Data Analysis

For visualization and comparisons, the percentage of choosing higher acoustic features by participant was calculated. For example, if one participant chose the latter one in the pair of L-L-L and L-L-M, it would get one score for being "choosing higher features" because medium-level formant dispersion is higher than the low-level one. Firstly, we used the proportion data to examine whether the choice pattern was significantly biased. Thus, we applied one-sample *t*-tests to compare the choice of higher parameters with the chance level (50%).

Secondly, the mixed-effect logistic regression was conducted using the *lme4* package in R (R Core Team, 2022) to compare the response patterns in raw data by the acoustic feature that was manipulated (F0, pitch range, formant dispersion), the step of comparisons (High-Medium: H-M, Medium-Low: M-L) and the education-level group (college and secondary school). The dependent variable is the responses to higher or lower features: If the participants chose the stimuli with higher features, their responses would be coded as 1; if they chose the lower features, their responses would be coded as 0. For example, if one participant chooses the L-L-M stimuli compared with L-L-L in one trial, his/her response will be coded as 1. The random intercepts were participants and trials.

# Results

As illustrated in Figure 2 and 3, the results showed that all participants had a clear preferential pattern of higher acoustic features of the perceived voice. The overall choices of higher features for both college students and secondary students were significantly higher than the chance level (college: $t(53) = 3.47, p = .001, 95\%CI[52.68; 60.05]$; secondary: $t(53) = 5.77, p < .001, 95\%CI[55.89; 62.17]$. However, their detailed response patterns were largely different from each other by condition. As shown in Table 1, college students tended to have a significantly larger proportion of choosing the Medium-wide pitch range than the chance level, while they did not show significant differences between the choices in other conditions and 50%. Similarly, the group of secondary students tended to regard wider pitch ranges as the features of being a good teacher, but the proportions of both H-M and M-L comparisons were significantly higher than the chance level. Besides, the secondary students also had a clear preference for higher acoustic cues in the H-M comparison of F0 and M-L comparison of Formant dispersion when they chose the features related to good teachers. Those suggest that the preference patterns of two age or education-level groups are potentially different from each other.

In addition, the students' preferences were further analyzed using the logistic regression model. The full model includes the acoustic feature that was manipulated (F0, Pitch Range, Formant Dispersion), the comparisons (High-Medium: H-M, Medium-Low: M-L), and the education-level group (college and secondary school). The model achieved an $R^2$ value of 0.158, indicating that only 15.8% of the variance in the dependent variable was accounted for by the fixed predictors and random effects. The only significant effect was the three-way interaction effect, $\chi^2(1) = 7.44, p = .024$. It indicates that the response patterns in each group were largely different from each other and the interactions of acoustic features and their comparison levels were also different. Thus, we conducted the post-hoc tests separately for each group. For the college students, there was no significant effect or changed features or comparisons (all $p$s $> .3$), suggesting the college students' choices were not significantly biased by the acoustic

Table 1: Summarized results of one-sample *t*-tests.

| Group | Acoustic features | Comparison | Mean (%) | *t*(8) | *p* | 95%*CI* |
|---|---|---|---|---|---|---|
| College | F0 | H-M | 54.86 | 0.96 | .367 | [43.15; 66.57] |
| | | M-L | 60.42 | 2.18 | .061 | [49.41; 71.42] |
| | Pitch range | H-M | 52.78 | 0.71 | .498 | [43.75; 61.80] |
| | | M-L | 63.89 | 3.16 | .013* | [53.76; 74.02] |
| | Formant dispersion | H-M | 53.47 | 0.69 | .508 | [41.92; 65.02] |
| | | M-L | 52.78 | 0.80 | .447 | [44.77; 60.78] |
| Secondary | F0 | H-M | 60.42 | 3.54 | .008** | [53.62; 67.21] |
| | | M-L | 54.17 | 0.97 | .360 | [44.26; 64.07] |
| | Pitch range | H-M | 59.72 | 2.40 | .043* | [50.38; 69.06] |
| | | M-L | 57.64 | 3.77 | .005** | [52.97; 62.31] |
| | Formant dispersion | H-M | 54.86 | 1.17 | .274 | [45.32; 64.40] |
| | | M-L | 67.36 | 4.07 | .004** | [57.52; 77.20] |

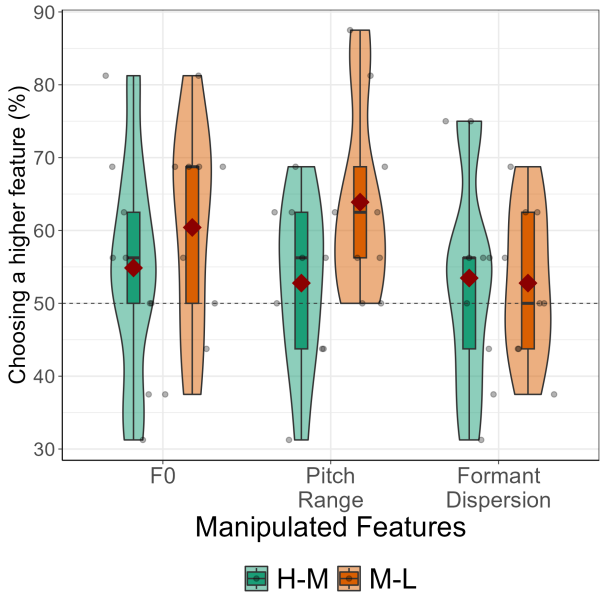


Figure 2: Preference for acoustic features of the college students. *Note*. the diamond-shaped points in dark red are the mean values.

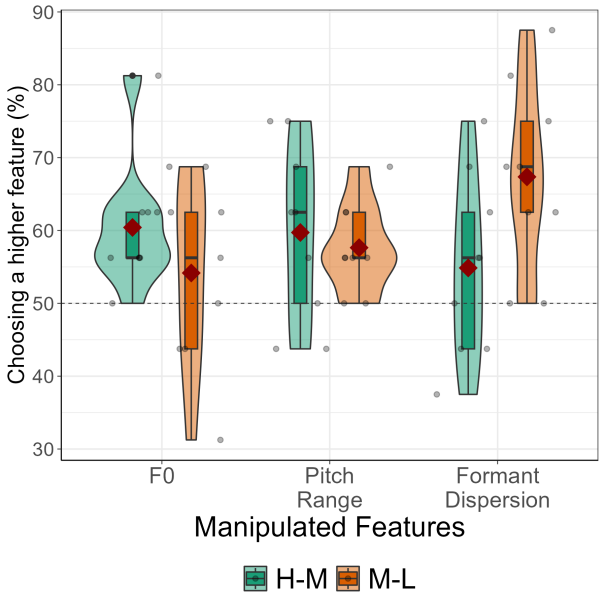


Figure 3: Preference for acoustic features of the secondary school students. *Note*. the diamond-shaped points in dark red are the mean values.

features and their manipulated levels.

Differently, for the secondary school students, their responses were significantly influenced by the interaction between the acoustic features and the comparisons H-M and M-L ($\chi^2(1) = 6.41, p = .041$). Thus, we conducted the following analyses of their interaction with the Tukey post-hoc tests and adjusted $p$ values, but there was no statistical significance between each comparison by acoustic feature and comparison level. Those indicate a subtle interactive pattern in the responses of the secondary school students.

## Discussion

The study aims to investigate how acoustic cues could affect students' perception of whether a teacher is good or not and whether that judgment could be affected by their education levels. Our results showed that both groups share a similar preference regarding the voices of a wider pitch range but only the secondary school students felt that a voice with a higher F0 and a higher level of formant dispersion might make a speaker sound like a good teacher. However, only the three-way interaction of the group, acoustic features, comparison levels, and the two-way interaction between acoustic features and comparison levels in the secondary school group were found to have significant effects on the listeners' per-

ception.

The divergent response patterns observed between the two age groups align with prior research suggesting that students' age and education level can shape their expectations of what constitutes a good teacher (Constantinou & Wijnen-Meijer, 2022; Wei et al., 2023). This is also consistent with studies indicating that age can influence personality perception and first impressions (Ewing, Austin, Diffin, & Grande, 2015). Notably, traits associated with trustworthiness are perceived at an earlier age rather than in adulthood (Oosterhof & Todorov, 2008). Future studies can be considered to narrow down to or select specific personal traits for a similar study setting. Apart from the relationship within the student group, this study also examines the generality of the pattern of students' vocal perception of a good teacher by conducting the perception task on adult participants. The result shows that the adult group tends to present several unique features regarding the perception of a good teacher, including the negative influence of a low level of formant dispersion which is significant. Except for the situation that students significantly tend to choose the voice with a wider pitch range than adults when the formant dispersion is fixed to a high level, the choices made by the students and adults indicate no statistical significance. Hence, it is rational to argue that regarding the vocal perception of a teacher's trait, age and cognitive development do not play an important part as expected in the research question.

Additionally, new features were found regarding the relationship between acoustic parameters and students' vocal perception of teacher's traits. It is found that from the student's perspective, the voices that were comparatively higher in fundamental frequency, wider in pitch range, and higher in the level of formant dispersion are more likely to be considered to be spoken by a good teacher. Among the three conditions, students' preference for voices with a wider pitch range shows statistical significance for both education-level groups. The finding agrees with the former research suggesting that a wider pitch range is considered more pleasant by the students and a higher pitch sounds more motivating to students (Gampel & Ferreira, 2017). For the study on formant dispersion, this study attempts to establish the relationship between a higher level of formant dispersion and the portrait of a good teacher, which furthered the understanding of students' vocal perception of a teacher's trait.

To explain the relationship between the acoustic parameters and students' preference for voice, it is worth noticing that the perceived portrait of the teacher's voice is linked closely with the potentially indicated traits of certain acoustic parameters. For example, the wider pitch range was commonly reported to be associated with the personality of being trustworthy and friendly (Gampel & Ferreira, 2017). Moreover, the higher fundamental frequency also indicated the trustworthiness of the speaker among male speakers (Tyler, 2015; Klofstad et al., 2012). Based on former studies regarding the linkage between acoustic cues and traits, it is rational to argue that from the students' perception, a teacher with a personality of being trustworthy and friendly is considered a better teacher than those with less trustworthiness. On the other hand, the acoustic parameter of formant dispersion, which reflected the length of the vocal tract, was reported to have a strong correlation with the traits related to dominance (Hughes et al., 2002). For example, the longer vocal tract results in lower formant dispersion, which is likely to be perceived as being dominated by the listeners. Hence, the student's preference in this study may suggest that a voice with less dominance, which is triggered by higher formant dispersion, is regarded as a better teacher according to the perception of student participants.

Combining aspects of the targeted acoustic parameters in our findings, the ideal personality of a good teacher could be possibly illustrated by the preferred condition of acoustic parameters: from the student's perspective, a teacher who is very trustworthy and not dominant is regarded as a good teacher from the uttered voice. Previous research suggested that whether students considered the teacher as being competent positively related to the student's perception of the teacher's personality (J. Jones, 1989). In fact, the relationship between a teacher's personality and teaching effectiveness was proved to exist as well (Polk, 2006). Specifically, it was found that the caring and trustworthiness personality was significantly correlated with the perceived expertise of teachers as rated by the students (Teven & Herring, 2005). In other words, the personality related to trustworthiness, kindness, and likeability tends to positively improve the teacher's perceived professionalism. Therefore, the current study supports the hypothesis that the teacher's personality would enhance teaching and learning by evidence-based the relation between the student's perception of specific vocal cues and the teacher's trait, which is a new perspective apart from the student's rating for the teacher and the teacher's self-report. Further studies that recruit more participants are anticipated to answer whether different groups affect perceptual preferences.

Although this study attempts to establish a solid relationship between acoustic parameters and students' perceptions of teachers' voices, there are still limitations that restrict the effectiveness of the findings. First, the $R^2$ value of our regression model was low (15.8%), and it may suggest that this model is not a good fit for our data. One possible reason is that we did not focus on describing the individual factors or biases in our dataset but our results potentially showed some individual patterns. Since this study focuses on the effect of acoustic parameters, further research can be conducted to explore the effects of individual differences (e.g., gender, language background, academic performance, cognitive ability, etc.), which may contribute to a better model in the future. Second, as one of our reviewers pointed out, our sample size and the diversity of word items may not be enough for a perception experiment. Thus, future studies are needed to achieve a larger sample size and more target words as well as conduct a power analysis before recruitment.

## Conclusion

This study attempts to establish the relationship between the acoustic features and the perceived traits of a teacher and further investigate the difference between adult and secondary students regarding the pattern of perception. To achieve such a goal, 16 secondary school students and 16 adults are recruited to conduct a psychological experiment that asks the participants to make judgments about traits and personalities between two voices. As a result, the study finds that students significantly prefer the voice with wider pitch range when choosing which one sounds like a good teacher, and tend to choose voice with higher fundamental frequency and formant dispersion. Several significant results related to the situation where certain condition is fixed are found as well. For the adult group, it is found that the general preference for the teacher's voice remains the same to a large extent, despite some special circumstances. The results indicate that the students tend to regard the voice that shows more trustworthiness and less dominance as being uttered by a good teacher, which could potentially benefit future teacher training and speech synthesis for educational purposes.

## References


Birkholz, P. (2013). Modeling consonant-vowel coarticulation for articulatory speech synthesis. *PloS one*, *8*(4), e60603.

Boersma, P. (2011). Praat: doing phonetics by computer [computer program].

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59.

Constantinou, C., & Wijnen-Meijer, M. (2022). Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. *BMC Medical Education*, *22*(1), 113.

Ewing, G., Austin, L., Diffin, J., & Grande, G. (2015). Developing a person-centred approach to carer assessment and support. *British journal of community nursing*, *20*(12), 580–584.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, *102*(2), 1213–1222.

Gampel, D., & Ferreira, L. P. (2017). How do adolescent students perceive aging teachers' voices? *Journal of Voice*, *31*(4), 512–e9.

Goy, H., Pichora-Fuller, M. K., & van Lieshout, P. (2016). Effects of age on speech and voice quality ratings. *The Journal of the Acoustical Society of America*, *139*(4), 1648–1659.

Hughes, S. M., Harrison, M. A., & Gallup Jr, G. G. (2002). The sound of symmetry: Voice as a marker of developmental instability. *Evolution and human behavior*, *23*(3), 173–180.

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *79*(1), 57–62.

Jones, J. (1989). Students' ratings of teacher personality and teaching competence. *Higher Education*, *18*(5), 551–558.

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. *PloS one*, *10*(8), e0133779.

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

Lee, W.-S., & Zee, E. (2003). Standard chinese (beijing). *Journal of the International Phonetic Association*, *33*(1), 109–112.

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language history questionnaire (lhq3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23*(5), 938–944.

McAleer, P., & Belin, P. (2018). The perception of personality traits from voices. In F. Sascha & P. Belin (Eds.), *The oxford handbook of voice perception.* Oxford University Press.

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'hello'? personality impressions from brief novel voices. *PloS one*, *9*(3), e90779.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior research methods*, *51*, 195–203.

Polk, J. A. (2006). Traits of effective teachers. *Arts education policy review*, *107*(4), 23–29.

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, *28*(5), 340–344.

R Core Team. (2022). R: a language and environment for statistical computing. *R Foundation for Statistical Computing*. Retrieved from `https://www.R-project.org/` doi: 10.1007/978-3-540-74686-7

Roach, P. (2004). British english: Received pronunciation. *Journal of the International Phonetic Association*, *34*(2), 239–245.

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

Schmidt, C. P., Andrews, M. L., & McCutcheon, J. W. (1998). An acoustical and perceptual analysis of the vocal behavior of classroom teachers. *Journal of Voice*, *12*(4), 434–443.

Schroeder, J., & Epley, N. (2015). The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate's appeal. *Psychological science*, *26*(6), 877–891.

Teven, J. J., & Herring, J. E. (2005). Teacher influence in the classroom: A preliminary investigation of perceived instructor power, credibility, and student satisfaction. *Communication Research Reports*, *22*(3), 235–246.

Tyler, J. C. (2015). Expanding and mapping the indexical field: Rising pitch, the uptalk stereotype, and perceptual variation. *Journal of English Linguistics*, *43*(4), 284–310.

Waller, S. S., Eriksson, M., & Sörqvist, P. (2015). Can you hear my age? influences of speech rate and speech spontaneity on estimation of speaker age. *Frontiers in psychology*, *6*, 144456.

Wei, X., Chow, M.-K., Huang, L., Huang, X., & Cheng, G. (2023). Teacher evaluation in primary and secondary schools: A systematic review of ssci journal publications from 2012 to 2022. *Sustainability*, *15*(9), 7280.

Xu, Y., Lee, A., Wu, W.-L., Liu, X., & Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PloS one*, *8*(4), e62397.