2018

Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type?

Nuttirudee Charoenruk
Chulalongkorn University, Bangkok, nuttirudee@cbs.chula.ac.th

Kristen Olson
University of Nebraska-Lincoln, kolson5@unl.edu

Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type?

Nuttirudee Charoenruk[1] and Kristen Olson[2]

1 Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok, Thailand

2 Department of Sociology, University of Nebraska–Lincoln, Lincoln, NE, USA

Corresponding author — Nuttirudee Charoenruk, Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Phayathai Road, Pathum Wan, Bangkok 10330, Thailand;
email nuttirudee@cbs.chula.ac.th

Abstract

As an aural mode, an interviewer's voice is an important part of telephone surveys. A speaker's voice can convey information about his or her trustworthiness and confidence, among other attributes. Consequently, respondents may perceive attributes of interviewers from their voices. We evaluated how five perceived attributes of interviewers (expertise, confidence, reliability, trustworthiness, and easiness to understand) are associated with interviewer voice characteristics (pitch, intonation, speech rate, and disfluencies) and whether this varies by types of question and interviewer characteristics. Overall, listeners perceived interviewers' attributes from their voices. Interviewers with higher pitched voices, moderate intonation, a faster pace, and fewer disfluencies were rated more positively, but these ratings differed somewhat for socially undesirable and complex questions. Reading questions at the recommended speech rate of two words per second leads to negative perceptions of interviewers.

Published in *Field Methods* 30:4 (2018), pp 312-328
DOI: 10.1177/1525822X18784500

Introduction

Survey interviewers' voices convey information to respondents about who they are and whether respondents can trust the interviewer with their information. Although prior research on interviewer voices has examined associations between vocal characteristics, perceived interviewer attributes, and unit nonresponse (e.g., Benkí et al. 2011; Groves et al. 2008; Oksenberg and Cannell 1988), whether listeners perceive interviewer attributes from their voices as they ask survey questions has not been investigated. Respondents tend to provide higher quality answers to interviewers perceived as being more credible (e.g., trustworthy, expert, reliable, and confident; Hovland et al. 1953) and easier to comprehend (Japec 2008). Furthermore, telephone survey practice typically instructs interviewers to ask questions with proper phrasing and inflection at an average pace of two words per second (wps; Cannell et al. 1981). This recommendation applies across all types of questions, but whether listeners perceive interviewers' voices similarly across socially desirable, undesirable, complex, and neutral questions has not been investigated. A lack of visual presence of an interviewer in a telephone survey may increase anonymity for respondents, potentially increasing data quality for sensitive questions (Jans 2010). If respondents perceive an interviewer's personality traits based on their voices, telephone interviews may not provide as much anonymity as expected.

Although acoustic properties affect perceptions of interviewers' voices during survey recruitment (e.g., Benkí et al. 2011), there are surprisingly few evaluations of acoustic properties of interviewer voices during administration of survey questions. This study extends previous paralinguistic research using voice characteristics to predict subjective evaluations of a speaker's attributes (e.g., Apple et al. 1979) to the asking of different types of survey questions. In particular, we empirically examine whether listeners perceive survey interviewers' attributes of easiness to understand, expertise, confidence, reliability, and trustworthiness from interviewer voice characteristics of pitch, intonation, speech rate, and disfluencies. We evaluate whether perceptions of interviewers' voices vary over interviewers' readings of socially desirable, socially undesirable, complex, and neutral survey questions, as well as for male versus female and experienced versus inexperienced interviewers.

The voice characteristics of pitch, intonation, speech rate, and disfluencies are important contributors to how a listener perceives a speaker's emotions and other attributes (Bänziger et al. 2014). High-pitched voices are perceived as less truthful, less reliable, less trustworthy, and more confident than lower pitched voices (Apple et al. 1979; Scherer et al. 1973; Van der Vaart et al. 2005). The association between pitch and perceptions of easiness to understand is inconsistent (Broome 2012; Van der Vaart et al. 2005). Intonation refers to a pattern of rising and falling pitch and to the patterns of stress in a language (Oksenberg and Cannell 1988). Voices with broad variations in pitch are perceived as being less reliable, less trustworthy, more confident, and easier to understand than voices with less variation (Scherer et al. 1973; Van der Vaart et al. 2005).

Rate of speech is the number of words spoken over the duration of speech (Webb 1969). Survey interviewers are instructed to speak at a rate of two wps (Cannell et al. 1981), slower than the speech rate in ordinary conversation (2.5–3.2 wps; Tauroza and Allison 1990). Rapid speech is perceived as more credible, more trustworthy, and expressing more confidence than slow speech (e.g., Apple et al. 1979; Broome 2012; Oksenberg and Cannell 1988; Smith and Shaffer 1995). However, faster speech may be harder to understand (Miller et al. 1976).

Disfluency is the parts of speech that are not words, such as fillers (ums and uhs; Conrad et al. 2008). Disfluencies can indicate that a speaker lacks confidence and is encountering difficulties (Bortfeld et al. 2001; Conrad et al. 2008; Ehlen et al. 2007). Highly disfluent speakers are also judged as less credible, less trustworthy, less easy to understand, and less expert than more fluent speakers (Apple et al. 1979; Ketrow 1990; Miller et al. 1976; Oksenberg and Cannell 1988).

We hypothesize that listeners will perceive interviewers who read questions with higher pitched voices and with more variable pitch (more intonation) as less reliable, less trustworthy, and more confident than those who read questions with lower pitched voices. Additionally, we expect that pitch will be associated with evaluations of expertise and easiness to understand, although we cannot anticipate a direction. We anticipate that more intonation will help ease of understanding and increase perceptions of expertise. We also expect listeners to perceive interviewers who read questions at a faster rate of speech and with fewer disfluencies as more trustworthy, more reliable, more expert, and more confident. We hypothesize that raters

will perceive faster speakers and those who speak with more disfluencies as less understandable.

We anticipate that the effects of voice characteristics on interviewers' attributes may not be linear (e.g., Conrad et al. 2008). For example, it may be easier to understand a speaker talking at a normal speech rate (around 3 wps) than someone speaking at either slower or faster speech rates. As such, we also examine nonlinear relationships between interviewers' voice characteristics and interviewers' attributes.

Interviewer characteristics such as gender and experience also may affect listeners' perceptions of interviewers' personal attributes. For example, males speak slightly faster than females (Yuan et al. 2006) and more experienced interviewers have shorter interview durations (Olson and Peytchev 2007). Consequently, female and inexperienced interviewers may be rated as less trustworthy, less reliable, less expert, and less confident but more understandable than male and experienced interviewers. Moreover, gendered stereotypes of voice characteristics indicate that females have higher pitched voices, greater intonation, somewhat slower speech rates, and use fewer fillers than males (Bortfeld et al. 2001; Kent and Read 2002). Additionally, Benkí et al. (2011) found that male interviewers with less vocal intonation had higher contact rates than male interviewers with more intonation, but the opposite effect occurred for female interviewers. Thus, we expect that voice attributes that deviate from gendered stereotypes may have adverse effects.

Different types of questions, such as those that are socially desirable, undesirable, or complex, may also affect how listeners perceive interviewers' voices. For example, because a high level of emotional arousal (e.g., fear and anxiety) is associated with an increase in voice pitch (Bachorowski 1999), interviewers may read socially undesirable questions with higher pitched voices than they use with other questions, and this may affect listeners' perceptions of attributes such as trustworthiness. Thus, we examine whether there are differences in perceptions across different types of questions. We now describe the data and method for evaluating these questions.

Data and Methods

Data for this study came from the Work and Leisure Today Survey conducted in 2013 by AbtSRBI (American Association of Public Opinion Research RR3 = 6.3%; American Association of Public Opinion Research 2016). It was a landline random-digit dial computer-assisted telephone survey administered by 22 interviewers with 450 completed interviews. To increase the stability of the analyses, interviewers who conducted fewer than 10 interviews were eliminated from the study (Olson and Peytchev 2007), resulting in 432 interviews conducted by 19 interviewers (nine females and 10 males) for the analysis.

Out of 24 candidate questions representative of the four categories— socially desirable, socially undesirable, complex, and neutral (not complex and not socially desirable or undesirable) in the questionnaire, we selected 12 questions (three from each of the four categories) based on the criteria that questions have sufficient variability in responses. The topics of these questions cover employment status and volunteer work, leisure activities such as exercise and Internet use, and substance use such as drinking alcohol (see Appendix Table A1). However, because of skipped questions, poor-quality voice files (n = 42), and voice files containing interruptions (n = 40), a sample of 4,689 voice files of interviewers reading individual questions were analyzed.

Voice Characteristics

We used Praat, a standard computer program used for linguistic analysis, for speech analysis (Boersma and Weenink 2010). Pitch and intonation were operationalized as the mean and the standard deviation of pitch over the period that an interviewer first read a survey question, respectively. The number of words spoken while reading a survey question divided by the number of seconds spent reading that question results in the speech rate (number of wps). Moreover, we used the number of fillers, including um, uh, ah, mm, hm, and oh, as the measure of disfluencies. The average interviewer pitch, intonation, speech rate, and number of fillers are 167.6 Hz (SD = 40.0), 41.7 Hz (SD = 19.9), 3.5 wps (SD = 1.0), and 0.2 fillers (SD = 0.5) per voice file, respectively.

Table 1. Definitions of Interviewer Attributes.

Attributes	*Definition*
Confidence	The extent to which the interviewer is self-assured and poised in conducting the interview
Easiness to understand	The extent to which the interviewer's voice is easy to understand
Reliability	The extent to which an interviewer says something that can be believed
Trustworthiness	The degree of confidence that an interviewer will ask valid survey questions and keep respondents' answers confidential
Expertise	The extent to which an interviewer is good at his or her job in asking survey questions

Subjective Ratings of Interviewers' Attributes

Six undergraduate students (four males and two females) coded interviewer attributes. Each coder listened to each of the voice files containing the first time an interviewer read each survey question to each respondent and rated five interviewer attributes (interviewers' confidence, easiness to understand, reliability, trustworthiness, and expertise) on seven-point scales from low (1) to high (7) (see **Table 1** for full definitions). We averaged the attribute ratings across the six coders.

Analyses

We examined the relationships between the interviewer voice characteristics and the interviewer attributes using two-level multilevel models to account for variation due to questions (i, level 1) and interviewers (j, level 2; Olson and Peytchev 2007). We also tested for a significant interviewer variance effect through a base model using a mixture of chi-square distributions (Rabe-Hesketh and Skrondal 2012). In general, we estimate:

$$\text{Rating}_{ij} = \gamma_{00} + \gamma_{01}\text{IChar}_{j} + \sum_{a=1}^{A}\gamma_{a0}\text{VoiceChar}_{aij} + \sum_{a=1}^{A}\gamma_{a1}\text{IChar}_{j} \times \text{VoiceChar}_{aij} + \sum_{b=A+1}^{B}\gamma_{b0}\text{QChar}_{bij} + \sum_{c=B+1}^{C}\gamma_{c0}\text{VoiceChar}_{ij} \times \text{QChar}_{ij} + U_{0j} + e_{ij}$$

with a random interviewer effect $U_{0j} \sim N(0, \tau^2)$

and residuals $e_{ij}^* \sim (0, \sigma^2)$, where:

Rating_{ij} = ratings of confidence, expertise, reliability, trustworthiness, and easiness of understand,

VoiceChar_{ij} = pitch, intonation, speech rate, and number of fillers, grand mean centered,

QChar_{ij} = neutral, complex, socially undesirable, and socially desirable questions, and

IChar_{j} = indicator variables for interviewer's gender and experience.

Nonlinear relationships between voice characteristics and rated interviewer attributes may exist (Conrad et al. 2008), and as such, we also examined the squared terms of pitch, intonation, speech rate, and number of fillers. Interviewers' demographic characteristics are dichotomous variables parameterized as female (= 1, 49%) versus male interviewers (= 0, 51%) and interviewers whose experience is one year or less (= 1, 25%) versus more than one year (= 0, 75%). Neutral questions are the reference group for types of questions.

We conducted all analyses using PROCMIXED with maximum likelihood estimation in SAS (Bell et al. 2013; Hoffman 2015). Additionally, the ESTIMATE statement in SAS was used to compute the expected rating of interviewer attributes at specific values of the independent variables to facilitate interpretation of nonlinear terms and interaction effects (Kiernan et al. 2011).

Results

The average ratings of interviewers' confidence, easiness to understand, reliability, trustworthiness, and expertise for each trait are 5.76 (SD = 0.53), 5.90 (SD = 0.49), 5.89 (SD = 0.43), 5.74 (SD = 0.50),

and 5.87 (SD = 0.50), respectively. We evaluated reliability across raters using the intraclass correlation coefficient (ICC). Using Munro's (2005) criteria of reliability, there was moderate reliability in confidence (ICC = .60) and expertise (ICC = .63) and low reliability in easiness to understand (ICC = .46), reliability (ICC = .43), and trustworthiness (ICC = .47). Using the average ratings, the chi-square tests for variance across interviewers is significant ($p < .01$) for all five rated interviewer attributes. That is, interviewers vary significantly in how raters perceive their confidence, reliability, trustworthiness, expertise, and easiness to understand.

Pitch

Pitch has a curvilinear relationship with rated confidence, reliability, trustworthiness, and expertise and has a linear relationship with rated easiness to understand (**Table 2**; standard errors in Appendix Table A2). Consistent with our hypotheses, interviewers who read a question with higher pitched voices are rated as being more confident, easier to understand, and having more expertise, but with a decelerating rate. However, counter to our hypotheses, raters view interviewers with higher pitched voices as more reliable and trustworthy.

Intonation

Intonation has an inverse U-shaped association with all five ratings of interviewer's attributes (**Figure 1**). Perceptions of interviewers are highest for moderate intonation (around 60 Hz), compared to lower or higher intonation, and decline notably when interviewers read questions with an intonation higher than 80 Hz, an intonation level observed for 5% of voice files in this study. That is, when interviewers read questions with an intonation higher than the typical range (intonation between 20 and 80 Hz, accounting for 95% of voice files), perceptions of interviewers decrease significantly.

Speech Rate

As expected, listeners rated a fast speech rate as more confident, more reliable, more trustworthy, and more expert compared to a slow speech rate. Additionally, we found an inverse U-shaped relationship

Table 2. Hierarchical Linear Model Coefficients Predicting Subjective Ratings of Interviewer Attributes by Voice Characteristics.

Variables	*Confidence*	*Easiness to Understand*	*Reliability*	*Trust-worthiness*	*Expertise*
Intercept	6.014**	6.143**	5.991**	5.782**	6.050**
Pitch	0.005**	0.0001**	0.002**	0.003**	0.003**
Pitch^2	_0.00002**	–0.00001	–0.00001**	–0.00002**	–0.00002**
Intonation	0.002**	0.003**	0.001**	0.002**	0.001**
Intonation^2	–0.00004**	–0.00001**	–0.00005**	–0.00004**	–0.00004**
Speech rate	0.096**	0.132**	0.108**	0.062**	0.119**
Speech rate^2	–0.013**	–0.089**	–0.029**	–0.020**	–0.034**
Filler	–0.120**	–0.118**	–0.090**	–0.028**	–0.124
Filler^2	0.002	0.026	0.009	0.000001	0.001
Interviewer experience					
1+ year	—	—	—	—	—
≤ 1 year	–0.144	0.119	–0.003	–0.053	–0.037
Interviewer gender					
Male	—	—	—	—	—
Female	–0.243	–0.069	–0.029	0.097	–0.071
Question type					
Neutral questions	—	—	—	—	—
Desirable questions	0.017	–0.069**	0.031	–0.002	–0.004
Complex questions	0.006	–0.242**	0.003	–0.017	–0.024
Undesirable questions	–0.105**	–0.169**	–0.068**	–0.122**	–0.122**
Variance components					
Interviewer	0.085**	0.052**	0.037**	0.068**	0.085**
Residual	0.148**	0.141**	0.124**	0.127**	0.138**
Model fit					
Akaike Information Criterion (AIC)	4,475.9	4,232.7	3,643.1	3,737.4	4,136.2
Bayesian information criterion (BIC)	4,491.1	4,247.8	3,658.2	3,752.5	4,151.3

N = 4,689

** p < .01

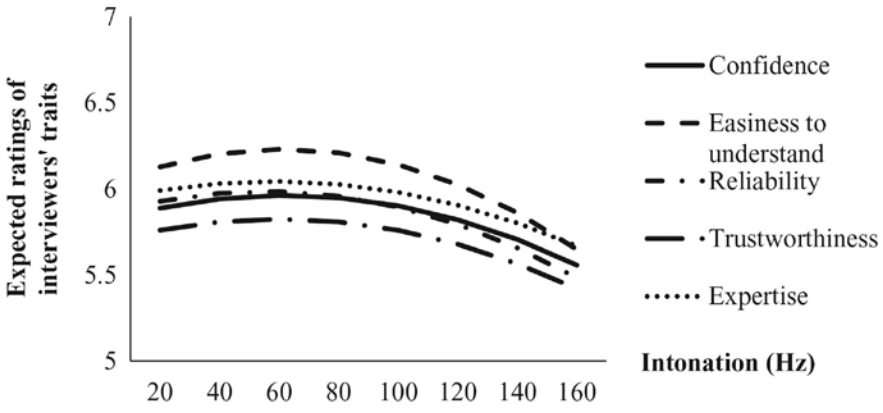


Figure 1. Expected ratings of interviewers' attributes by intonation

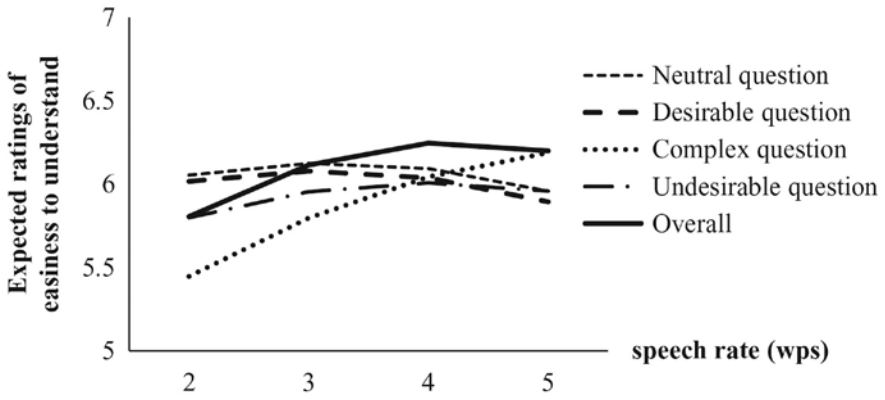


Figure 2. Expected ratings of easiness to understand by speech rate and question type.

between expected ratings of easiness to understand and speech rate with the highest rating of easiness to understand at 4 wps (overall line; **Figure 2**).

Fillers

As expected, interviewers who asked questions with more fillers were rated as being less confident, less easy to understand, less reliable, and less trustworthy than those who asked questions with fewer fillers.

Interviewer Voices and Interviewer Demographic Characteristics

Ratings of interviewers do not differ by interviewer experience, but they do vary by interviewer sex. Female interviewers were rated as less confident in asking survey questions than male interviewers; however, the effect is just marginally significant. Additionally, interviewer sex moderates some effects of intonation, speech rate, and fillers on rated attributes of interviewers. The full moderation models are in Appendix Table A4.

Intonation and Interviewer Sex

Interviewer sex moderates the effects of intonation on rated confidence (female × intonation coefficient = .005, $p < .01$), rated reliability

(female × intonation coefficient = .002, p = .02), rated trustworthiness (female × intonation coefficient = .003, p < .01), and rated expertise (female × intonation coefficient = .003, p < .01; Appendix Table A3). Male interviewers were rated as being more confident, more reliable, and having more expertise than female interviews at low intonation levels, but the opposite was found at high intonation levels (Appendix Figure A1). Male voices receive the highest expected ratings of confidence, reliability, trustworthiness, and expertise at an intonation of 40 Hz. For female voices, the highest expected ratings of confidence, reliability, trustworthiness, and expertise are at intonations ranging between 60 and 80 Hz. Thus, lower ratings on each trait occur for male voices with high intonation and for female voices with low intonation, consistent with gendered expectations of voices.

Speech Rate and Interviewer Sex

The differences for speech rate on rated easiness to understand and expertise by gender are modest (female × speech rate coefficient = .047, p < .01 for easiness to understand, female × speech rate coefficient = 0.27, p = 0.03 for expertise; Online Table A4). As speech rate interests, male interviewers were rated as slightly more expert and easier to understand than female interviewers. However, at a speech rate of 4 and 5 wps, male voices were rated as slightly less easy to understand than female voices.

Filler and Interviewer Sex

The negative effect of fillers on rated trustworthiness was significant only for female interviewers (female × speech rate coefficient = −0.069, p < 0.01). Female interviewers who read questions wither fewer fillers were rated as more trustworthy than those who read questions with more fillers.

Interviewer Voices and Question Characteristics

Ratings of interviewers' attributes vary by question type. Raters judged interviewers reading socially undesirable questions as less confident (coefficient = −.09), less easy to understand (coefficient

= −.128), less reliable (coefficient = −.065), less trustworthy (coefficient = −.118), and less expert (coefficient = −.119) than interviewers reading neutral questions ($p < .01$; Appendix Table A4). Furthermore, interviewers reading socially desirable questions were rated as more reliable than those reading neutral questions (coefficient = .035, $p = .02$; Appendix Table A4), but less easy to understand (coefficient = −.051, $p < .01$; Appendix Table A4). Additionally, compared to neutral questions, raters evaluated interviewers as less easy to understand for *complex questions* (coefficient = −.189, $p < .01$; Appendix Table A4). In addition to the main effects of question types on rated interviewer attributes, question types moderate the effects of intonation, speech rate, and fillers on rated interviewers' attributes, especially rated confidence and easiness to understand.

Intonation and Question Types

The effect of intonation on rated easiness to understand differs between complex and neutral questions (intonation × complex coefficient = .003, $p < .01$; Appendix Table A4). Neutral questions generally are rated as easier to understand than complex questions, but the difference between the two types of questions decreases as intonation increases.

Speech Rate and Question Types

The effects of speech rate on rated confidence and easiness to understand differ by question type. As speech rate increases, rated confidence increases with decelerating rate; the effect was strongest in socially desirable questions. In contrast, the *direction* of the effects of speech rate on rated easiness to understand significantly *differs* by question types. An inverse U-shaped relationship exists for neutral, desirable, and undesirable questions (Figure 2), with the highest expected rating of easiness to understand at 3 wps for neutral and socially desirable questions and at 4 wps for socially undesirable questions. For complex questions, however, speech rate is positively associated with rated easiness to understand, such that complex questions asked more quickly are rated as easier to understand.

Fillers and Question Types

The negative effect of fillers on rated easiness to understand was stronger in socially undesirable questions than in neutral questions (coefficient = –.136, $p < .01$; Appendix Table A4). Socially undesirable survey questions read with more fillers are rated as less easy to understand than neutral questions.

Discussion

Overall, we found that listeners rated interviewers as having more positive interviewer attributes when interviewers read questions with higher pitched voices, moderate intonation, a speech rate faster than the recommended pace of 2 wps, and fewer fillers. Counter our hypotheses, raters view interviewers with higher pitched voices and higher intonation as more reliable and trustworthy, possibly because low-pitched voices and low-pitch variation are associated with undesirable attributes such as dishonesty and dominance, whereas high-pitched voices and high intonation are associated with positive attributes such as friendliness (Boehme 2014).

Following gendered stereotypes of voices, perceptions of interviewer attributes from intonation and speech rate vary by interviewer gender. As females have higher intonation and speak more slowly than males (Bortfeld et al. 2001; Kent and Read 2002; Yuan et al. 2006), perceptions of interviewer attributes for male interviewers were more positive at *lower* intonation levels and faster speech rates. In contrast, perceptions of interviewer's attributes for female interviewers were more positive at *higher* intonation levels.

This study is the first to show that question types moderate how listeners evaluate a question's easiness to understand. Overall, the highest rating of easiness to understand was found when interviewers read questions with a speech rate of 4 wps, especially for complex questions. Counter to common survey practice recommendations (Cannell et al. 1981), for complex questions, listeners perceived speech faster than 2 wps and with more intonation as easier to understand. Listeners may find it difficult to hold the relevant information about complex questions in working memory when interviewer read questions slowly.

The included variables differ in their explanatory power across the models (Appendix Table A5). Although only 4.9% of the interviewer-level variance in the easiness to understand ratings is explained by the included covariates, between 15.3% (expertise) and 38% (trustworthiness) of the interviewer-related variance is explained for the other ratings. At the question level, the included covariates explain about 10% of the variance for all ratings except for easiness to understand, where 30.5% of the question-level variance is explained. Furthermore, speech rate has the largest explanatory value of the question-level variance across the perceptions of interviewers (Appendix Table A3). Thus, interviewer voice characteristics explain more variability *across questions* in how easy they are to understand but more variability *across interviewers* in how the interviewers themselves are perceived for the other characteristics.

This study has limitations. Undergraduate coders may not represent perceptions of interviewers at large. Listeners at different age cohorts judge voices based on different criteria; young people use different standards for those who are from their generation than older adults (Ketrow 1990). Thus, perceptions of interviewers by older adults may vary somewhat, especially for the effects of fast speech on perceptions and overall perceptions of easiness to understand, where older adults may have more difficulty understanding speech than younger adults (e.g., Gordon-Salant et al. 2007). We also could not directly study variability across coders because only six undergraduate coders were used here, although the average of raters' evaluations tends to be quite reliable (Olson 2010). With only two female coders, it is not possible in these data to evaluate gender differences in perceptions of interviewer voices; future research should examine how different attributes of coders affect perceptions of interviewers. The study is also limited to a single set of interviewers in one geographic area, and we did not measure the interviewer's accent, a vocal feature also shown to affect perceptions of speakers (e.g., Bradac and Wisegarver 1984).

Ultimately, perceptions of interviewer voices matter if the perceptions also affect data quality. For instance, respondents may provide better data to interviewers who they perceive as more credible or easier to understand (Japec 2008). Thus, the negative perceptions of interviewers with more vocal intonation or more speech fillers may result in lower data quality from those interviewers. Additionally,

interviewers who speak quickly and are difficult to understand may connote to respondents that it is not necessary to take time to give accurate and complete answers, also lowering data quality. It may be that the modest effects of some of the vocal characteristics observed here on perceptions have only limited effects on data quality. Future research will examine how actual and perceived voice characteristics affect answers to these survey questions.

Results in this study suggest that survey organizations should instruct interviewers to read questions with a speech rate faster than 2 wps (especially for complex questions) to be perceived more positively. This is notably different from the recommendation to read survey questions at a pace of two wps (Cannell et al. 1981). Slightly faster question reading also may reduce survey interview length, potentially lowering survey costs.

Declaration — The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. They disclosed receipt of the following financial support and thank the National Science Foundation (NSF SES-1356985 and NSF SES-1132015) for supporting this study.

Supplemental material follows the References.

References

American Association for Public Opinion Research. 2016. Standard definitions: Final dispositions of case codes and outcome rates for surveys. http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (accessed June 17, 2017).

Apple, W., L. A. Streeter, and R. M. Krauss. 1979. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology* 37: 715–27.

Bachorowski, J.-A. 1999. Vocal expression and perception of emotion. *Current Directions in Psychological Science* 8:53–57.

Bänziger, T., S. Patel, and K. R. Scherer. 2014. The role of perceived voice and speech characteristics in vocal emotion communication. *Journal of Nonverbal Behavior* 38:31–52.

Bell, B. A., M. Ene, W. Smiley, and J. A. Schoeneberger. 2013. A multilevel model primer using SAS PROC MIXED. http://support.sas.com/resources/papers/proceedings13/433-2013.pdf (accessed June 17, 2017).

Benkí, J., J. Broome, F. Conrad, R. Groves, and F. Kreuter. 2011. Effects of speech rate, pitch, and pausing on survey participation decisions. *Paper presented at the American Association for Public Opinion Research Annual Meeting*, Phoenix, AZ.

Boehme, B. 2014. How trustworthy is your voice? The effects of voice manipulation on the perceived trustworthiness of novel speakers. Unpublished PhD diss., University of Glasgow, Glasgow, UK.

Boersma, P., and D. Weenink. 2010. fPgraat: Doing phonetics by computer. http://www.praat.org (accessed October 4, 2014).

Bortfeld, H., S. D. Leon, J. E. Bloom, M. F. Schober, and S. E. Brennan. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44:123–47.

Bradac, J. J., and R. Wisegarver. 1984. Ascribed status, lexical diversity, and accent: Determinants of perceived status, solidarity, and control of speech style. *Journal of Language and Social Psychology* 3:239–55.

Broome, J. S. 2012. Vocal characteristics, speech, and behavior of telephone interviewers. Unpublished PhD diss., University of Michigan, Ann Arbor, MI.

Cannell, C. F., P. V. Miller, and L. Oksenberg. 1981. Research on interviewing techniques. *Sociological Methodology* 12:389–437.

Conrad, F. G., M. F. Schober, and W. Dijkstra. 2008. Cues of communication difficulty in telephone interviews. In *Advances in telephone survey methodology*, eds. J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, and R. L. Sangster, 212–30. New York: Wiley.

Ehlen, P., M. F. Schober, and F. G. Conrad. 2007. Modeling speech disfluency to predict conceptual misalignment in speech survey interfaces. *Discourse Processes* 44:245–65.

Gordon-Salant, S., P. J. Fitzgibbons, and S. A. Friedman. 2007. Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners. *Journal of Speech, Language, and Hearing Research* 50: 1181–93.

Groves, R. M., B. C. O'Hare, D. Gould-Smith, J. Benki, and P. Maher. 2008. Telephone interviewer voice characteristics and the survey participation decision. In *Advances in Telephone Survey Methodology*, eds. J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, and R. L. Sangster, 385–400. New York: Wiley.

Hoffman, L. 2015. *Longitudinal analysis: Modeling within-person fluctuation and change*. New York: Routledge.

Hovland, C. I., K. J. Irving, and H. K. Harold. 1953. *Communication and persuasion*. New Haven, CT: Yale University Press.

Jans, M. E. 2010. Verbal paradata and survey error: Respondent speech voice, and question-answering behavior can predict income item nonresponse. Doctoral dissertation. University of Michigan, Ann Arbor, MI. http://hdl.handle.net/2027.42/75932 (accessed June 5, 2017).

Japec, L. 2008. Interviewer error and interviewer burden. In *Advances in telephone survey methodology*, eds. J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, and R. L. Sangster, 185–211. New York: Wiley.

Kent, R. D., and C. Read. 2002. *Acoustic analysis of speech*, 2nd ed. Albany, NY: Thomson Learning.

Ketrow, S. M. 1990. Attributes of a telemarketer's voice and persuasiveness. A review and synthesis of the literature. *Journal of Direct Marketing* 4:7–21.

Kiernan, K., R. Tobias, P. Gibbs, and J. Tai. 2011. CONTRAST and ESTIMATE statements made easy: The LSMESTIMTE statement. https://support.sas.com/resources/papers/proceedings11/351-2011.pdf (accessed June 17, 2017).

Miller, N., G. Maruyama, R. J. Beaber, and K. Valone. 1976. Speed of speech and persuasion. *Journal of Personality and Social Psychology* 34:615.

Munro, B. 2005. *Statistical methods for health care research*. London: Lippincott Williams & Wilkins.

Oksenberg, L., and C. Cannell. 1988. Effects of interviewer vocal characteristics on nonresponse. In *Telephone Survey Methodology*, eds. R. M Groves, P. B. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nichols II, and J. Waksberg, 257–69. New York: Wiley.

Olson, K. 2010. An examination of questionnaire evaluation by expert reviewers. *Field Methods* 22:295–318.

Olson, K., and A. Peytchev. 2007. Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly* 71:273–86.

Rabe-Hesketh, S., and A. Skrondal. 2012. *Multilevel and longitudinal modeling using Stata*. College Station, TX: Stata Press.

Scherer, K. R., H. London, and J. J. Wolf. 1973. The voice of confidence: Paralinguistic cues and audience evaluation. *Journal of Research in Personality* 7: 31–34.

Smith, S. M., and D. R. Shaffer. 1995. Speed of speech and persuasion: Evidence for multiple effects. *Personality and Social Psychology Bulletin* 21:1051–60.

Tauroza, S., and D. Allison. 1990. Speech rates in British English. *Applied Linguistics* 11:90–105.

Van der Vaart, W., Y. Ongena, A. Hoogendoorn, and W. Dijkstra. 2005. Do interviewers' voice characteristics influence cooperation rates in telephone surveys? *International Journal of Public Opinion Research* 18:488–99.

Webb, J. T. 1969. Subject speech rates as a function of interviewer behaviour. *Language and Speech* 12:54–67.

Yuan, J., M. Liberman, and C. Cieri. 2006. Towards an integrated understanding of speaking rate in conversation. In *Proceedings of INTERSPEECH 2006*, 541–44. Grenoble: International Speech Communication Association. http://languagelog.ldc.upenn.edu/myl/ldc/llog/icslp06_final.pdf (accessed June 20, 2017).

Appendix Materials

(included for review purposes)

Appendix Table A1. Questions by question types

Question characteristics	**Question and response options**
Complex questions Complex questions can create problems for the cognitive response process. For example, the question contains an unfamiliar word.	**Q13E** (On a scale from 1 to 5 where 5 means you enjoy the activity completely and 1 means you do not enjoy the activity at all, please tell me how much you enjoy) Kaninhop? 5 Enjoy completely 4 Enjoy a lot 3 Enjoy somewhat 2 Enjoy a little 1 Do not enjoy at all 8 DK 9 REF **Q19** On a typical day, how many minutes do you spend on a computer? DEFINITION (ON THE SAM E SCREEN AS THE QUESTION): There are 1,440 minutes in a 24-hour day. ☐ 8888 DK 9999 REF

	Q20 In the past week, how many email messages, if any, have you written or received? 88888 DK 99999 REF
S ocially undesirable questions Socially undesirable questions are questions to which respondents tend to underreport their answers.	**Q5** Have you ever been fired from a job? 1 Yes 2 No 8 DK 9 REF **Q21C** (Thinking about the past seven days, how many times did you)- drink alcohol? [INTERVIEWER: We are interested in the total number of drinks.] 88888 DK 99999 REF

Question characteristics	**Questions and response options**
	Q21D (Thinking about the past seven days, how many times did you) have sex?  88888 DK 99999 REF
S ocially desirable questions Socially desirable questions are questions to which respondents tend to overreport their answers.	**Q8** We are interested in volunteer activities for which people are not paid, except perhaps expenses. We only want you to include volunteer activities that you did through or for an organization, even if you only did them once in a while. In the last 12 months, that is since July of last year (2012), have you done any volunteer activities through or for an organization? [PROBE: IF HAVE NOT VOLUNTEERED, ASK:] Sometimes people don't think of activities they do infrequently or activities they do for children's schools or youth organizations as volunteer activities. Since July of last year, have you done any of these types of volunteer activities?

	1 Yes 2 No 8 DK 9 REF **Q13A** On a scale from 1 to 5 where 5 means you enjoy the activity completely and 1 means you do not enjoy the activity at all, please tell me how much you enjoy the following leisure activities. First, how about reading? 5 Enjoy completely 4 Enjoy a lot 3 Enjoy somewhat 2 Enjoy a little 1 Do not enjoy at all 8 DK 9 REF **Q21F** (Thinking about the past seven days, how many times did you) read a book, magazine, or newspaper  88888 DK 99999 REF

Question characteristics	Questions and response options
Neutral questions Neutral questions are not complex and not socially desirable/undesirable questions.	**Q2** Compared to 10 years ago (2013), do you think people have more leisure time, less leisure time, or about the same amount? 1 More 2 Same amount 3 Less 8 DK 9 REF **Q13D** (On a scale from 1 to 5 where 5 means you enjoy the activity completely and 1 means you do not enjoy the activity at all, please tell me how much you enjoy) fishing or hunting. 5 Enjoy completely 4 Enjoy a lot 3 Enjoy somewhat 2 Enjoy a little 1 Do not enjoy at all 8 DK 9 REF **Q21A** People do a number of different types of activities for leisure. Thinking about the past seven days, how many times did you use the internet?

	 88888 DK 99999 REF

Appendix Table A2. Hierarchical linear model coefficients and standard errors predicting subjective ratings of interviewer attributes by voice characteristics

	Confidence			**Easiness to Understand**			**Reliability**			**Trustworthiness**			**Expertise**		
	Coef.	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***	
Intercept	6.014	*0.10*	**	6.143	*0.08*	**	5.991	*0.10*	**	5.782	*0.09*	**	6.050	*0.10*	**
Pitch	0.005	*0.0004*	**	0.0001	*0.0004*	**	0.002	*0.0004*	**	0.003	*0.0004*	**	0.003	*0.0004*	**
Pitch^2	−0.00002	*0.000004*	**	−0.00001	*0.000004*		−0.00001	*0.000004*	**	−0.00002	*0.000004*	**	−0.00002	*0.000004*	**
Intonation	0.002	*0.0005*	**	0.003	*0.005*	**	0.001	*0.0004*	**	0.002	*0.0004*	**	0.001	*0.0005*	**
Intonation^2	−0.00004	*0.00001*	**	−0.0001	*0.00001*	**	−0.00005	*0.00001*	**	−0.00004	*0.00001*	**	−0.00004	*0.00001*	**
Speech rate	0.096	*0.007*	**	0.132	*0.01*	**	0.108	*0.007*	**	0.062	*0.01*	**	0.119	*0.007*	**
Speech rate^2	−0.013	*0.004*	**	−0.089	*0.004*	**	−0.029	*0.004*	**	−0.020	*0.004*	**	−0.034	*0.004*	**
Filler	-0.120	*0.03*	**	−0.118	*0.03*	**	−0.090	*0.026*	**	−0.028	*0.03*	**	−0.124	*0.03*	
Filler^2	0.002	*0.01*		0.026	*0.014*		0.009	*0.01*		0.000001	*0.014*		0.001	*0.014*	
Iwer exp.≤1 year	−0.144	*0.15*		0.119	*0.12*		−0.003	*0.10*		−0.053	*0.10*		−0.037	*0.15*	
Female Iwer	−0.243	*0.14*		−0.069	*0.11*		−0.029	*0.09*		0.097	*0.12*		−0.071	*0.14*	
Desirable Qs	0.017	*0.016*		−0.069	*0.16*	**	0.031	*0.01*		−0.002	*0.01*		−0.004	*0.016*	
Complex Qs	0.006	*0.017*		−0.242	*0.17*	**	0.003	*0.02*		−0.017	*0.02*		−0.024	*0.017*	

Undesirable														
Qs	−0.105	*0.018*	**	−0.169	*0.17*	**	−0.068	*0.02*	**	−0.122	*0.02*	**	−0.122	*0.017*	**
Variance components														
Interviewer	0.085	*0.03*	**	0.052	*0.02*	**	0.037	*0.01*	**	0.068	*0.02*	**	0.085	*0.03*	**
Residual	0.148	*0.003*	**	0.141	*0.003*	**	0.124	*0.002*	**	0.127	*0.003*	**	0.138	*0.003*	**
Model fit														
AIC	4,475.9			4,232.7			3,643.1			3,737.4			4,136.2		
BIC	4,491.1			4,247.8			3,658.2			3,752.5			4,151.3		

Note: N=4,689; **p<0.01, *p<0.05

Appendix Table A3. Variance components, standard errors, and percent of explained variance predicting subjective ratings of interviewer attributes for empty model with no covariates and model in Appendix Table 2 with voice characteristics, interviewer characteristics, and question characteristics

	Confidence			**Easiness to Understand**			**Reliability**			**Trustworthiness**			**Expertise**		
	Coef.	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***	
Variance components															
Model with all interviewer vocal characteristics															
Interviewer	0.08493	*0.03*	**	0.05184	*0.02*	**	0.03723	*0.01*	**	0.06803	*0.02*	**	0.08494	*0.03*	**
Question	0.1481	*0.03*	**	0.1408	*0.003*	**	0.1243	*0.002*	**	0.1265	*0.003*	**	0.1377	*0.003*	**
Excluding pitch and pitch^2															

Interviewer	0.09576	*0.03*	**	0.04976	*0.02*	**	0.03960	*0.01*	**	0.07497	*0.02*	**	0.08736	*0.03*	**
Question	0.1515	*0.003*	**	0.1412	*0.003*	**	0.1251	*0.003*	**	0.1280	*0.003*	**	0.1391	*0.003*	**
Excluding intonation and intonation[2]															
Interviewer	0.08867	*0.03*	**	0.05489	*0.02*	**	0.03838	*0.013*	**	0.07497	*0.02*	**	0.08693	*0.03*	**
Question	0.1486	*0.003*	**	0.1420	*0.003*	**	0.1248	*0.003*	**	0.1280	*0.003*	**	0.1380	*0.003*	**
Excluding speech rate and speech rate[2]															
Interviewer	0.09108	*0.03*	**	0.04435	*0.01*	**	0.04121	*0.01*	**	0.07326	*0.02*	**	0.08951	*0.03*	**
Question	0.1539	*0.003*	**	0.1673	*0.003*	**	0.1332	*0.003*	**	0.1297	*0.003*	**	0.1487	*0.003*	**
Excluding fillers and fillers[2]															
Interviewer	0.08258	*0.03*	**	0.05359	*0.02*	**	0.03637	*0.01*	**	0.06687	*0.02*	**	0.08338	*0.027*	**

Question	0.1505	*0.003*	**	0.1418	*0.003*	**	0.1253	*0.003*	**	0.1267	*0.003*	**	0.1403	*0.003*	**
% Explained Variance															
By pitch & pitch^2															
Interviewer	11.3%			-4.2%			6.0%			9.3%			2.8%		
Question	2.2%			0.3%			0.6%			1.2%			1.0%		
By intonation & intonation^2															
Interviewer	4.2%			5.6%			3.0%			9.3%			2.3%		
Question	0.3%			0.8%			0.4%			1.2%			0.2%		
By speech rate & speech rate^2															

Interviewer	6.8%	-16.9%	9.7%	7.1%	5.1%
Question	3.8%	15.8%	6.7%	2.5%	7.4%
By fillers & fillers^2					
Interviewer	-2.8%	3.3%	-2.4%	-1.7%	-1.9%
Question	1.6%	0.7%	0.8%	0.2%	1.9%

Note: N=4689; **p<0.01, *p<0.05

Appendix Table A4. Hierarchical linear model coefficients and standard errors predicting subjective ratings of interviewer attributes by voice characteristics, interviewer characteristics, and question characteristics

	Confidence			Easiness to Understand			Reliability			Trustworthiness			Expertise		
	Coef.	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***	
Intercept	6.027	*0.10*	**	6.070	*0.08*	**	5.991	*0.07*	**	5.778	*0.087*	**	6.049	*0.10*	**
Pitch	0.005	*0.0004*	**	0.0006	*0.0004*		0.002	*0.0004*	**	0.003	*0.0004*	**	0.003	*0.0004*	**
Pitch^2	−0.00003	*0.000004*	**				−0.00001	*0.000004*	**	−0.00002	*0.0000004*	**	−0.00002	*0.000004*	**
Intnation	0.0004	*0.0006*		0.001	*0.0007*	*	0.0007	*0.0005*		0.0005	*0.0005*		0.0003	*0.0006*	
Intonation^2	−0.00005	*0.000012*	**	−0.0001	*0.00001*	**	−0.0001	*0.00001*	**	−0.00004	*0.00001*	**	−0.00004	*0.00001*	**
Speech rate	0.135	*0.018*	**	−0.032	*0.02*		0.108	*0.006*	**	0.062	*0.007*	**	0.107	*0.009*	**
Speech rate^2	−0.017	*0.006*	**	−0.051	*0.005*	**	−0.029	*0.004*	**	−0.020	*0.004*	**	−0.033	*0.004*	**
Filler	−0.118	*0.01*	**	−0.046	*0.02*	*	−0.075	*0.01*	**	−0.0009	*0.016*		−0.123	*0.01*	**
Iwer exp. ≤1 year	−0.129	*0.15*		0.103	*0.12*		−0.029	*0.10*		0.058	*0.13*		−0.031	*0.15*	
Female Iwer	−0.276	*0.13*		−0.034	*0.11*		−0.039	*0.09*		0.093	*0.12*		−0.085	*0.14*	
Desirable Qs	0.022	*0.02*		-−0.051	*0.02*	**	0.035	*0.01*	*	0.002	*0.01*		0.0003	*0.02*	
Complex Qs	−0.010	*0.02*		−0.189	*0.02*	**	0.005	*0.02*		−0.015	*0.02*		−0.024	*0.02*	
Undesirable Qs	−0.090	*0.02*	**	−0.128	*0.02*	**	−0.065	*0.02*	**	−0.118	*0.02*	**	−0.119	*0.02*	**

Voice * Iwer chars

Intonation*female Iwer	0.005	*0.0008*	**				0.002	*0.0008*	*	0.003	*0.0008*	**	0.003	*0.001*	**
Speech rate*female Iwer				0.047	*0.01*	**							0.027	*0.013*	*
Filler*female Iwer										−0.069	*0.02*	**			

Voice * question chars

Intonation*desirable Qs				−0.0001	*0.0008*	
Intonation*complex Qs				0.003	*0.0008*	**
Intonation*undesirable Qs				0.0002	*0.0008*	
Speech rate*desirable Qs	0.054	*0.03*	*	-0.009	*0.02*	
Speech rate*complex Qs	-0.068	*0.02*	**	0.279	*0.02*	**
Speech rate*undesirable Qs	-0.054	*0.02*	*	0.085	*0.02*	**
Filler*desirable Qs				0.004	*0.03*	
Filler*comple x Qs				-0.033	*0.04*	
Filler*undesirable Qs				-0.136	*0.04*	**

Variance components															
Interviewer	0.080	*0.03*	**	0.05	*0.02*	**	0.04	*0.01*	**	0.064	*0.02*	**	0.083	*0.03*	**
Residual	0.146	*0.003*	**	0.132	*0.003*	**	0.124	*0.003*	**	0.126	*0.003*	**	0.137	*0.002*	**
Model fit															
AIC	4,416.6			3,952.7			3,637.8			3,717.6			4,122.9		
BIC	4,434.5			3,975.4			3,652.9			3,733.7			4,139.0		

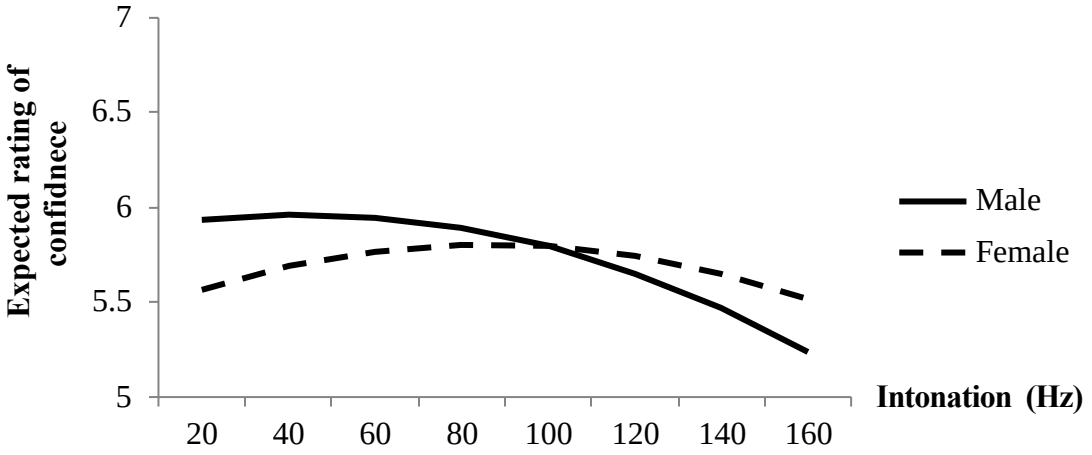
Note: N=4,689; **p<0.01, *p<0.05

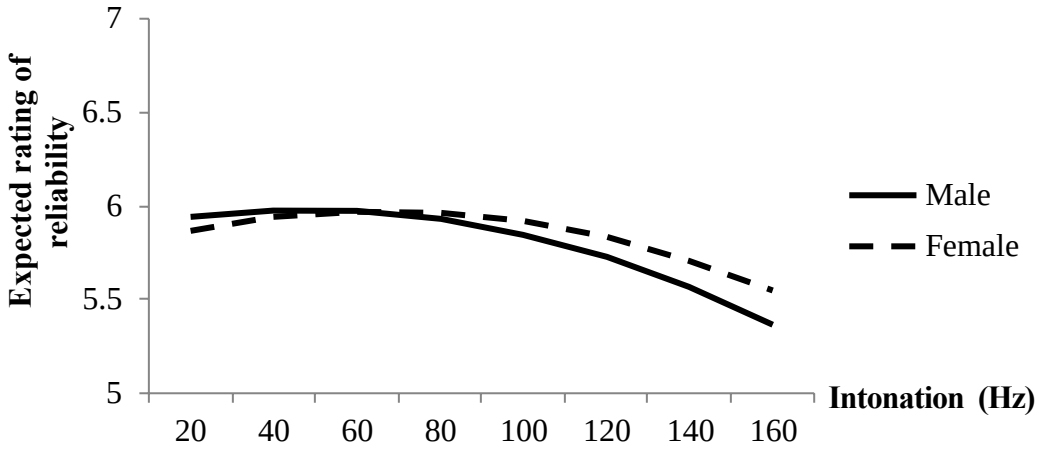
Appendix Table A5. Variance components, standard errors, and percent of explained variance predicting subjective ratings of interviewer attributes for empty model with no covariates and full model with voice characteristics, interviewer characteristics, and question characteristics

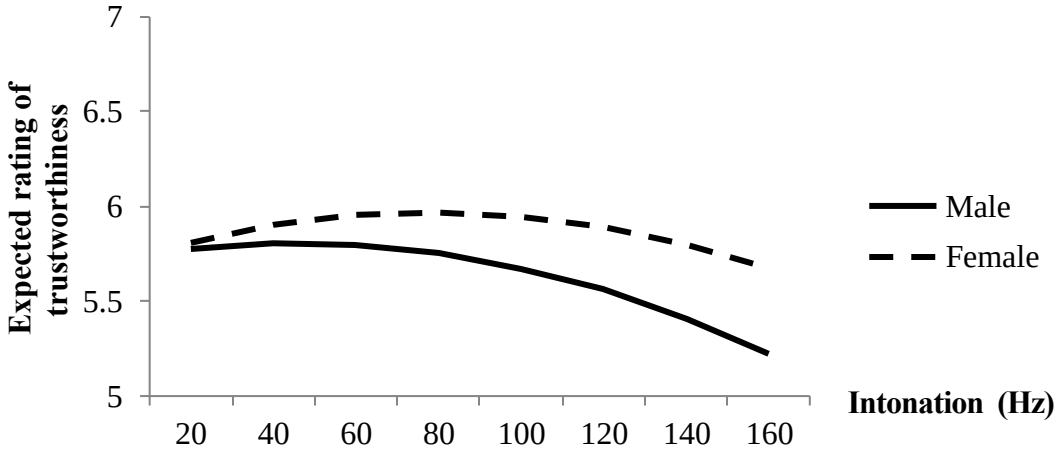
	Confidence			Easiness to Understand			Reliability			Trustworthiness			Expertise		
	Coef.	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***		**Coef.**	***SE***	
Variance components															
Base Model															
Interviewer	0.116	*0.04*	**	0.052	*0.02*	**	0.048	*0.02*	**	0.104	*0.03*	**	0.098	*0.03*	**
Question	0.162	*0.003*	**	0.190	*0.004*	**	0.137	*0.003*	**	0.134	*0.003*	**	0.155	*0.003*	**
Full Model															
Interviewer	0.080	*0.03*	**	0.049	*0.02*	**	0.036	*0.01*	**	0.064	*0.02*	**	0.083	*0.03*	**
Question	0.146	*0.003*	**	0.132	*0.003*	**	0.124	*0.003*	**	0.126	*0.003*	**	0.137	*0.002*	**

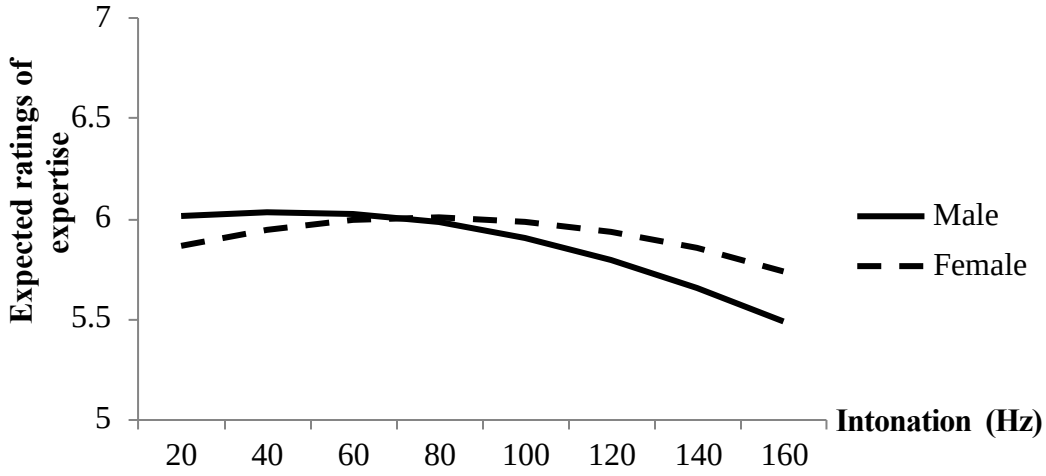
% Explained					
Variance					
Interviewer	30.9%	4.9%	24.6%	38.0%	15.3%
Question	9.7%	30.5%	9.3%	6.0%	11.4%

Note: N=4689; **p<0.01, *p<0.05


Expected rating of confidnece
7
6.5
6
5.5
5
20
40
60
80
100
120
140
160
Intonation (Hz)
Male
Female


Expected rating of reliability
7
6.5
6
5.5
5
20
40
60
80
100
120
140
160
Intonation (Hz)
Male
Female


Expected rating of trustworthiness
7
6.5
6
5.5
5
20
40
60
80
100
120
140
160
Intonation (Hz)
Male
Female



Appendix Figure A1. Expected ratings of confidence, reliability, trustworthiness, and expertise by intonation and interviewer sex