INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Speech Likability and Personality-based Social Relations: A Round-Robin Analysis over Communication Channels

Laura Fernández Gallardo, Benjamin Weiss

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Germany
(laura.fernandezgallardo|benjamin.weiss)@tu-berlin.de

Abstract

The Social Relations Model is well-known for analyses of interpersonal attraction. As a novelty in this paper, the model is applied to assess different effects on likability ratings from speech only. A group of 30 unacquainted participants is considered in our experiment. Their voices were recorded and transmitted through communication channels, and ratings of speech likability and speaker personality were then collected from the same individuals following a round-robin approach. This setup enabled us to detect the influence of participants' personality and of narrowband and wideband speech on the sources of variance according to the Social Relations Model. An analysis of acoustic correlates of speech likability has also been conducted, which shows differences in the relevance of speech features and in the description of likability ratings depending on the speech bandwidth.

Index Terms: speaker likability, speaker personality, Social Relations Model, communication channels

1. Introduction

Previous research has shown that individuals' social characteristics such as likability and personality can be detected from the speech signal. Personality has been defined in [1] as *"a psychological construct aimed at explaining the wide variety of human behaviors in terms of a few, stable and measurable individual characteristics"*, and voice likability—or voice pleasantness—is regarded as the question of *"how much we like a speaker based on the sound of her/his voice and manner of speaking"* [2]. Despite having applied advanced feature extraction and machine learning techniques for the automatic detection of these two social characteristics, only a modest accuracy of 60–70% has been reached on the binary classification task [3]. This highlights the need of further research establishing relationships between human perceptions and speech features, which should lead to machines reaching the human performance in the recognition of speaker characteristics.

In order to delve deeper into human perceptions of speaker likability and personality, this paper presents a round-robin experimental design. This design, based on the Social Relations Model (SRM) [4], consists on a group of persons (at least four participants are required) who mutually rate each another in terms of interpersonal perceptions. In our work, we employ the SRM as a statistical approach to analyze mutual perceptions of speaker likability and personality. The interpersonal ratings can then be decomposed into three independent sources of variance: (a) perceiver variance, which accounts for the variability of ratings introduced by the raters or annotators, (b) target variance, accounting for the variability of the given ratings within the persons being rated, and (c) relationship variance, that is, the variance in the person's behavior toward another individual in particular, considering all participants pairs. By applying this approach, personality ratings at hand can be used to explain specific likability variance components. These are typically discarded by averaging, as current research in paralinguistics focuses on target variance only.

The SRM has already been used to study liking for unacquainted people [5], yet not taking speech characteristics into account. As a recent exception, a SRM analysis to investigate the effect of personality on interpersonal attraction considered audio-visual impressions subjectively rated by external annotators [6]. The authors collected 73 x 72 interpersonal judgments of likability in their round-robin design. Among other results, the study revealed that extroverted persons and persons with strong, non-nervous, gender typical, and friendly voices were more liked. The present study has a similar aim, yet only speech material (no video) is collected and then presented to the group of participants for labeling. Our round-robin design involves 30 participants, thus yielding 30 x 29 interpersonal perceptions of likability and of personality from voice only. We then examine the personality characteristics of perceivers and targets which determine the SRM sources of variance. Only zero acquaintance scenarios are considered.

So far, no study has reported the use of the SRM for speech-based interpersonal relations, to the best of our knowledge. Other investigations of speech likability and personality perceptions have recruited external annotators, not found within the speakers, who provided their ratings on a given scale [2, 7, 8, 9]. Based on these ratings, previous studies analyzed different correlations between perceptions and acoustic parameters. In [10], male speakers with low F0 and female speakers with energy spread over the spectrum and lower 3rd central moment were perceived as more likable, as well as speakers with higher articulation rate and lower spectral center of gravity. A number of investigations have reported that low F0 results in more pleasant and attractive voices [11, 12]. Also, increased F0 dynamics and fluency correlated positively with the impressions of good speakers in [13]. Pronunciation also plays an important role in voice likability. Listeners seem to prefer to hear voices with their own accent [14] or with no accent [15].

The second focus of the present work is the study of the effect of communication channel bandwidth on the SRM sources of variance. Commonly, we unintentionally try to compile a picture of our interlocutors over the phone from their voice characteristics, which may determine our decisions and attitudes toward the speakers. Notwithstanding the current rapid deployment of digital communications, the influence of different telephone transmissions on voice likability perceptions has been overlooked. The parameters of the communication channels in telephone datasets have been treated as a "black box"

http://dx.doi.org/10.21437/Interspeech.2016-459

and no comparison between different transmission conditions has been found to be published.

While narrowband (NB, 300–3,400 Hz) is still predominant in most of today's communication networks, efforts have been made towards motivating the transition to an extended frequency band, namely wideband (WB, 50–7,000 Hz). Benefits of WB over NB channels have been shown for speech quality perception [16], human and automatic speaker recognition [17], intelligibility [17], and speech recognition [18]. This paper includes a preliminary analysis of acoustic correlates of likability from NB and from WB voices, which seems to indicate an advantage of the extended channel bandwidth for the description of speech likability ratings.

2. Round-robin design

A group of 30 persons (15f, 15m) with a mean age of 27.2 years (range: 20–34) and German as mother tongue participated in our round-robin experiment. As a requisite for our study, we only recruited speakers of standard High German dialect, from whom no marked accent could be perceived by another native German speaker. These 30 persons were invited on two separated days with an interval of approximately two weeks: first to a recording session and then to perform a listening test in which to rate the likability and personality of each other participant from their voice. Except for two people who were a couple (1f, 1m), the rest of participants indicated not to be acquainted with the other test persons and they never saw each another. Each of the sessions took one hour to complete. They were compensated for their participation.

2.1. Speech data collection

The recording session was divided into two parts. In the first part, the participants were asked to read words, digits, sentences, and turns from four different dialogs. The dialogs simulated telephone calls held with a female German speaker who assisted the recordings, and always played the role of a contact person or agent. The recorded speech, from the client's side, included different inquiries about some information: to a health insurance company, a mobile telecommunications company, a car rental company, and a real estate agency, for each dialog, respectively. The speakers were asked to read the exact given text as naturally as possible, yet without emotions or exaggerated friendliness.

While the first recording part comprised prescribed texts, spontaneous speech was elicited in the second part. The participants were asked to silently read a short story (of 17 sentences) and to utter a 30-second summary of its content. Afterwards, four spontaneous telephone dialogs were held: renting a car, ordering a pizza, ordering a book from the library, and making an appointment at the doctor's. These dialogs follow the scenarios known as Short Conversation Tests, found in the International Telecommunication Union (ITU)-T Rec. P.805. Again, the participants played the client's role and the recording assistant the agent's role. A summary of the recorded speech contents is given in Table 1.

The speech was recorded employing the RME Fireface UCX Audio Interface with 48 kHz sampling frequency and 32-bit quantization and using the software Cubase 4. The microphone used was AKG C 414B-XLS (frequency range 20–20,000 Hz) and was mounted on a boom stand in an acoustically-isolated room (Dimensions: 2.75 m x 2.53 m x 2.10 m, RT60 = 0.08 s at 2 kHz).

Table 1: *Recorded speech. 30 speakers (15f, 15m), high-quality microphone, 48 kHz sampling frequency.*

Prescribed speech	Words, digits, sentences Dialog 1: Health insurance Dialog 2: Mobile phone rate plan* Dialog 3: Car rental—inquiry Dialog 4: Real estate agency
Spontaneous speech	Short story summary Dialog 5: Car rental—booking Dialog 6: Pizza* Dialog 7: Book from the library Dialog 8: Doctor's appointment

* parts selected for the listening test

2.2. Preparation of the speech stimuli

Excerpts were extracted from the described high-quality recordings and prepared for the listening test, which involved two sections. In the first section, listeners rated speaker likability by listening to speech transmitted through a NB or a WB channel. The same fixed sentence was selected from all speakers: "Ich würde auf die SMS gern verzichten und meine Frei-Minuten dafür erhöhen" (In English: "I would like to give up the SMS and increase my free minutes in return"). This segment was extracted from the dialog with the telecommunications company, in which all speakers uttered the same text. The sentences had a mean duration of 4.4 s and a standard deviation of 0.3 s.

These excerpts were transmitted through a NB and a WB communication channel. The speech was first level-equalized 26 dB below the overload of the digital system (-26 dBov) by applying the ITU-T Rec. P.56. Afterwards, they were bandwidth-filtered complying with the ITU-T G.712 and P.341 for NB and WB, respectively. The coding-decoding processes G.711 at 64 kbit/s for NB and G.722 at 64 kbit/s for WB were then applied using standard ITU tools for transmission channel simulation. This resulted in 60 versions of the same sentence (30 speakers and two channel conditions).

In the second section of the listening test, differently, the participants rated speaker likability and personality by listening to the spontaneous dialog in which the speakers ordered a pizza. The agent's speech was removed (there was never an overlap) and silences of 0.7 s duration were inserted between turns. Part of the dialog in which contact details were exchanged was also removed. The final mean duration of the dialog turns, including silences, was 19.5 s with standard deviation of 4.3 s. These signals were downsampled to 44.1 kHz via an anti-aliasing low-pass FIR filter and leveled to -26 dBov. No further processing was applied.

2.3. Listening test

The audio stimuli described in the previous subsection were presented to the participants in the listening test. Importantly, they did not listen to their own voice. Hence, they rated the 29 other participants (14 persons of the same gender and 15 persons of the opposite gender).

Before the listening test started, the participants completed the BFI-10 questionnaire referring to their own personality. Then, in the first part of the listening test, the NB and the WB sentence stimuli were presented to the participants in random order. They were asked to rate the voice likability after listening to each stimulus, and they could only listen to it once. The male voices were first presented, followed by the set of female voices. In total, 58 stimuli were presented in this part, corresponding to 29 speakers and two channel conditions. The listeners hence rated the same speaker's sentence twice (in NB and in WB), yet they were not informed that the same voices would appear twice, or that different telephone qualities were included within the audio stimuli. In the second part of the listening test, the spontaneous dialog turns corresponding to the 29 other speakers were presented, also randomly and separated into male and female sets. The participants were asked to rate the voice likability and each question of the Big-Five inventory of 10 items (BFI-10) [19] referring to the speaker's personality. In this second section, they could listen to each stimulus as many times as they wished.

All stimuli were rated in terms of likability by clicking on a slider with the antonyms "sympathisch" and "unsympathisch" (in English: "likable" and "non-likable") at its ends. For the personality assessments, sliders with the labels "trifft berhaupt nicht zu" and "trifft voll und ganz zu" (in English: "disagree strongly" and "agree strongly") were employed to rate every BFI-10 questionnaire item. No pre-defined indicator, marks, numeric values or other labels was shown on the sliders.

For the whole test, the listeners were instructed not to consider the stimuli content or the quality of the recordings. Instead, they were asked to base their answers on the sound of the voice heard and manner of speaking. Short pauses were inserted every 15 minutes, approximately, in order to avoid listeners' fatigue and loss of focus. The test was administered in a quiet office room using a laptop with a standard sound card and the closed headphones AKG K601 (frequency response 12–39,500 Hz) with diotic listening.

3. Results

The conducted analyses and respective findings are reported in the next three subsections. First, the effects of the participants' personality traits on the three SRM sources of variance of the likability perceptions (perceiver, target and relationship variances) are investigated. Afterwards, the effects of the NB and WB communication channels on likability and on the SRM sources of variance are assessed. Finally, using the NB and WB speech, acoustic correlates of speaker likability are detected and the influence of channel bandwidth discussed.

The R package TripleR was employed for the computation of the SRM variance components [20]. The likability and personality ratings mutually given by the couple of participants who know each other (ratings of Participant 006 to Participant 015 and vice-versa) were removed in all our analyses.

3.1. Speaker personality and SRM variances

The effects of personality traits on liking were studied with data from the second part of the listening test, in which speech from the pizza dialog was employed. The SRM variance components on the likability variable are presented in Table 2. The distribution of variance is comparable to other studies (e.g. [6]), with a similar low amount of perceiver and target variance and a high amount of relationship variance. The perceiver effect reflects the participant's tendency to rate positively or negatively, whereas the target effect measures how likable a speaker is on average (the most common data analyzed). The remaining variance in the ratings comprises the effect of individual dyadic relationships. In the following, correlation analyses conducted separately for the three effects are presented.

Table 2: *Relative variance components of liking (clean speech).*

Variance component	standardized	t.value
Perceiver	.148	3.384**
Target	.150	3.389**
Relationship	.702	20.043***

$**p < .01$; $***p < .001$

Table 3: *Personality effects on the SRM variance components.*

Trait	Perceiver	Target	Relationship
Extroversion	.042	.419*	.012
Agreeableness	.309	.512**	.124***
Conscientiousness	.110	.002	.067.
Neuroticism	.243	-.193	.152***
Openness	-.065	.325	-.163***

$.p < .10$; $*p < .05$; $**p < .01$; $***p < .001$

3.1.1. Perceiver effects: "What is the personality of people who like others?"

Correlating the Big-Five self reports with perceiver variance yields no significant result (cf. Table 3), although an effect of the Agreeableness trait was hypothesized. Collecting additional personality profiles on the participants from close friends might have improved the validity and reliability of the personality traits. However, such a study was not conducted and the personality attributions from the other participants cannot be considered as they do not know each other. These are, in turn, relevant for the target effects.

3.1.2. Target effects: "Does attributed personality affect likability ratings?"

The ratings of the Big-Five traits (z-normalized for better comparability) were correlated with target variance (cf. Table 3). Two effects have been found. Assumed extrovert and agreeable speakers are rated more positively. This is a well known effect for acquainted people [21, 22]. However, for unacquainted people, the study in [6] reports only a positive effect of Extroversion on target effects, whereas Agreeableness was not correlated. Still, there is sufficient evidence suggesting a positive effect of perceived benevolence on liking and social attractiveness for the acoustic domain and unacquainted people, e.g. [23].

3.1.3. Relationship effects: The remaining variance after controlling for target and perceiver

Specifically, considering a pair of participants A–B, the relationship variance accounts for the extent to which A likes B controlling for A's general tendency toward liking others and B's general tendency to be liked by others. According to the attraction theory [24], the most prominent factors to explain individual relationships are similarity, reciprocal likability, and proximity. Reciprocal likability, however, cannot be analyzed for relationship variance with our data because the participants never interacted. Therefore, any signal of liking, interest, or agreeableness would only affect the target variance. This might partly explain our correlation of Agreeableness–target variance. Also, with unacquainted participants recruited from a single city, only similarity remains as factor to be studied here.

Similarity between each dyad was obtained as the absolute difference between a perceiver's self-rating and his/her individual rating to his/her partner for each of the Big-Five traits, mul-

Table 4: *Relative standardized variance components of liking.*

Variance component	Narrowband	Wideband
Perceiver	.215***	.143**
Target	.087**	.160***
Relationship	.698***	.697***

$^{**}p < .01$; $^{***}p < .001$

tiplied by -1 and z-normalized. In contrast to [6], which did not find a similarity effect for personality but only for preferences (clothing, subculture), a significant similarity effect for Agreeableness and Neuroticism can be observed from our results (cf. Table 3). This indicates that pairs of individuals who are close in these two personality traits tend to like each other's voices to a greater extent, controlling by their tendencies of "being a liker" and of "being liked" by others. A plausible explanation for the dissimilarity effect for Openness is currently missing.

3.2. Effects of communication channels

The influence of NB and WB communication channels on voice likability were analyzed employing the ratings of the first part of the listening test. Considering a scale from 0 to 100, where greater numbers would represent higher perceived likability, the average rating was 45.47 and 53.30 for NB and WB speech, respectively (this is equivalent to the averaged target's effects). A two-sample t-test shows that the WB mean is significantly higher than the NB mean; $t(1718) = -7.354, p < .001$, revealing listeners' predilection for WB voices.

The SRM variance components computed from the NB and WB ratings separately are presented in Table 4. Again, a great value has been obtained for the relationship variance. Interestingly, large perceiver variance has been found with respect to the target variance in NB, while in WB both perceiver and target variance are similar. Two deductions can then be made: First, perceivers differ in "being a liker" to a greater extent in NB than in WB. This implies more variation among perceivers considering the average of their likability ratings of NB voices compared to WB. In other words, there is a wider range of positive and negative liking tendencies in NB, probably due to the perceivers' different degrees of tolerance to the NB distortion. And second, because targets differ in "being liked" more in WB than in NB, the differences between non-likable and likable voices appear to be better detected with the extended bandwidth.

3.3. Acoustic correlates of likability in NB and in WB

In order to detect correlates of target likability ratings, a small set of speech features has been considered based on previous experimentation and findings [10, 15]. The following features were extracted from the NB and WB speech using Praat [25], and z-normalized: Intensity (median and range); fundamental frequency (F0, median and range); speech duration; harmonic-to-noise ratio (HNR); center of gravity (CoG); and alpha ratio (calculated as the difference of energy in dB between the [1–5 kHz] and the [0.05–1 kHz] frequency bands).

Linear regression models have been estimated for NB and for WB data separately, using gender as a factor variable, and are presented in Table 5. It can be observed that the WB model contains a greater number of significant predictors than the NB model, and that the regression ($R^2 = .726$) is closer to be significant ($p = .084$ as opposed to $p = .742$). A step-wise feature selection yielded a significant WB model whereas no significance was found for the NB regression. In the models of Table 5, however, all predictors are included for a comparison between the two channels. The CoG and the alpha ratio are the most significant predictors of likability of male voices in WB, indicating a preference for darker voices [26] and for glottal adduction [27]. The effect of CoG is in concordance with [10].

Table 5: *Regression models for target likability in NB and in WB with acoustic predictors. NB model: $R^2 = .468$; $p = .742$. WB model: $R^2 = .726$; $p = .084$.*

Coefficients	Estimate (NB)	Estimate (WB)
females:Intensity-median	-.230	.665
males:Intensity-median	-.183	.220
females:Intensity-range	-.038	.512*
males:Intensity-range	.052	.010
females:F0-median	.440	-.023
males:F0-median	-.384	.581
females:F0-range	-.268	-.248
males:F0-range	-.213	-1.139.
females:duration	-.514	-.343
males:duration	.289	1.067*
females:HNR	-.261	-.660
males:HNR	-.138	-.716
females:CoG	.052	.306
males:CoG	-.602	-1.618**
females:alpha ratio	.324	-.301
males:alpha ratio	.646	1.211**

$.p < .10$; $^{*}p < .05$; $^{**}p < .01$

4. Conclusions

By means of a round-robin experiment, this paper has investigated effects of individuals' personality and of communication channels on speech likability assessments. Our analysis of the SRM sources of variance has shown that persons which are perceived as extroverted and agreeable are also rated with a higher likability. In addition, people similar in agreeableness and neuroticism tend to rate each other's voice likability more positively. WB voices, with respect to NB, are significantly higher rated in terms of likability on average, lead to lower variance among perceivers' rating tendencies, and allow listeners to better distinguish between non-likable and likable speech. These findings may motivate the further deployment of speech-based applications using WB- instead of NB-transmitted speech.

According to our examination of acoustic correlates of likability, perceivers' ratings can be better described in WB than in NB using our reduced set of features. It remains the question of whether our results generalize to a greater set of speakers and speech features, and using more sophisticated regression methods. Such a study is left for future work. Our ongoing research aims at recording a greater number of speakers and assessing their personality and likability from speech. The high-quality recordings will permit the subsequent transmission through NB and WB channels. The intention is to gather enough data to study the effects of bandwidth and codecs on the automatic prediction of these speaker characteristics—so far, only NB datasets have been released for this task.

5. Acknowledgments

This work has been supported by the German Research Foundation (DFG, Grant FE 1603/1-1 to Laura Fernández Gallardo and Grant WE 5050/1-1 to Benjamin Weiss).

6. References

[1] A. Vinciarelli and G. Mohammadi, "A Survey of Personality Computing," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 273–291, 2014.

[2] F. Burkhardt, B. Schuller, B. Weiss, and F. Weninger, "'Would you Buy a Car From Me?' – On the Likability of Telephone Voices," in *Interspeech*, 2011, pp. 1557–1560.

[3] B. Schuller, S. Steidl, A. Batliner, E. Nöth, A. Vinciarelli, F. Burkhardt, R. van Son, F. Weninger, F. Eyben, T. Bocklet, G. Mohammadi, and B. Weiss, "A Survey on Perceived Speaker Traits: Personality, Likability, Pathology, and the First Challenge," *Computer Speech & Language*, vol. 29, no. 1, pp. 100–131, 2015.

[4] D. A. Kenny, *Interpersonal Perception: A Social Relations Analysis*. New York, U. S.: Guilford Press, 1994.

[5] A. Chapdelaine, D. A. Kenny, and K. M. Lafontana, "Matchmaker, Matchmaker, Can You Make Me a Match? Predicting Liking Between Two Unacquainted Persons," *Journal of Personality and Social Psychology*, vol. 67, no. 1, pp. 83–91, 1994.

[6] M. D. Back, S. C. Schmukle, and B. Egloff, "A Closer Look at First Sight: Social Relations Lens Model Analysis of Personality and Interpersonal Attraction at Zero Acquaintance," *European Journal of Personality*, vol. 25, no. 3, pp. 225–238, 2011.

[7] A. Gravano, R. Levitan, L. Willson, v. Beňuš, J. Hirschberg, and A. Nenkova, "Acoustic and Prosodic Correlates of Social Behavior," in *Interspeech*, 2011, pp. 97–100.

[8] G. Mohammadi and A. Vinciarelli, "Automatic Personality Perception: Prediction of Trait Attribution Based on Prosodic Features," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 273–284, 2012.

[9] A. V. Ivanov, G. Riccardi, A. J. Sporka, and J. Franc, "Recognition of Personality Traits from Human Spoken Conversations," in *Interspeech*, 2011, pp. 1549–1552.

[10] B. Weiss and F. Burkhardt, "Voice Attributes Affecting Likability Perception," in *Interspeech*, 2010, pp. 1934–1937.

[11] M. Zuckerman, K. Miyake, and C. S. Elkin, "Effects of Attractiveness and Maturity of Face and Voice on Interpersonal Impressions," *Journal of Research in Personality*, vol. 29, no. 2, pp. 253–272, 1995.

[12] M. Babel, G. McGuire, and J. King, "Towards a More Nuanced View of Vocal Attractiveness," *PLoS One*, vol. 9, no. 2, pp. 1–10, 2014.

[13] E. Strangert and J. Gustafson, "What makes a good speaker? Subject Ratings, Acoustic Measurements and Perceptual Evaluations," in *Interspeech*, 2008, pp. 1688–1691.

[14] N. Dahlbäck, Q. Wang, C. Nass, and J. Alwin, "Similarity is More Important than Expertise: Accent Effects in Speech Interfaces," in *Conference on Human Factors in Computing Systems*, 2007, pp. 1553–1556.

[15] B. Weiss and F. Burkhardt, "Is 'Not Bad' Good Enough? Aspects of Unknown Voices' Likability," in *Interspeech*, 2012, pp. 510–513.

[16] S. Möller, A. Raake, N. Kitawaki, A. Takahashi, and M. Wältermann, "Impairment Factor Framework for Wideband Speech Codecs," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 1969–1976, 2006.

[17] L. Fernández Gallardo, *Human and Automatic Speaker Recognition over Telecommunication Channels*. Singapore: Springer, 2016.

[18] A. V. Ramana, L. Parayitam, and M. S. Pala, "Investigation of Automatic Speech Recognition Performance and Mean Opinion Scores for Different Standard Speech and Audio Codecs," *IETE Journal of Research*, vol. 58, no. 2, pp. 121–129, 2012.

[19] B. Rammstedt and O. P. John, "Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German," *Journal of Research in Personality*, vol. 41, no. 1, pp. 203–212, 2007.

[20] F. D. Schönbrodt, M. D. Back, and S. C. Schmukle, "TripleR: An R Package for Social Relations Analyses Based on Round Robin Designs," *Behavior Research Methods*, vol. 44, no. 2, pp. 455–470, 2012.

[21] D. van der Linden, R. H. Scholte, A. H. Cillessen, J. te Nijenhuis, and E. Segers, "Classroom Ratings of Likeability and Popularity are Related to the Big Five and the General Factor of Personality," *Journal of Research in Personality*, vol. 44, no. 5, pp. 669–672, 2010.

[22] J. Wortman and D. Wood, "The Personality Traits of Liked People," *Journal of Research in Personality*, vol. 45, no. 6, pp. 519–528, 2011.

[23] R. L. J. Street and R. M. Brady, "Speech Rate Acceptance Ranges as a Function of Evaluative Domain, Listener Speech Rate, and Communication Context," *Communication Monographs*, vol. 49, no. 4, pp. 290–308, 1982.

[24] E. Aronson, T. D. Wilson, and R. M. Akert, *Social Psychology*. Prentice Hall, 2009.

[25] P. Boersma and D. Weenink, *Praat: doing Phonetics by Computer [Computer program]*, Version 6.0.09, retrieved January 2016 from http://www.praat.org/, 2016.

[26] J. M. Grey and J. W. Gordon, "Perceptual Effects of Spectral Modifications on Musical Timbres," *The Journal of the Acoustical Society of America*, vol. 63, no. 5, pp. 1493–1500, 1978.

[27] J. Gauffin and J. Sundberg, "Data on the Glottal Voice Source Behavior in Vowel Production," *Transmission Laboratory, Quarterly Progress and Status Report*, vol. 21, no. 2–3, pp. 61–70, 1980.