Journal of Experimental Psychology: General

Acting in Action: Prosodic Analysis of Character Portrayal During Acting

Matthew Berry and Steven Brown

Online First Publication, May 13, 2019. http://dx.doi.org/10.1037/xge0000624

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

0096-3445/19/$12.00

http://dx.doi.org/10.1037/xge0000624

Acting in Action: Prosodic Analysis of Character Portrayal During Acting

Matthew Berry and Steven Brown
McMaster University

During the process of acting, actors have to embody the characters that they are portraying by changing their vocal and gestural features to match standard conceptions of the characters. In this experimental study of acting, we had professional actors portray a series of stock characters (e.g., king, bully, lover), which were organized according to a predictive scheme based on the 2 orthogonal personality dimensions of assertiveness and cooperativeness. We measured 12 prosodic features of the actors' vocal productions, as related to pitch, loudness, timbre, and duration/timing. The results showed a significant effect of character assertiveness on all 12 vocal parameters, but a weaker effect of cooperativeness on fewer vocal parameters. These findings comprise the first experimental analysis of vocal gesturing during character portrayal in actors and demonstrate that actors reliably manipulate prosodic cues in a contrastive manner to differentiate characters based on their personality traits.

Keywords: acting, gesture, performance, characters, prosody

Supplemental materials: http://dx.doi.org/10.1037/xge0000624.supp

In *The Republic*, Plato (380 BCE/1968) made a distinction between two basic modes of storytelling. In "diegesis", a story is told using the voice of a narrator, whereas in "mimesis", a story is told using the voices of the characters, as produced by actors who portray these characters in theatrical (and more recently cinematic) works. For the ancient Greeks, theater was considered to be an "imitative" art (Kristeller, 1951) because theater works were seen as mimetic depictions of people and events that could occur in the real world. An actor creates a portrayal of a person who could actually exist, and the challenge of the actor is to produce as realistic and compelling a representation of that person as is possible, not least when the character is quite different from the actor.

How an actor accomplishes this has been the subject of debate for well over two thousand years. Both historically and cross-culturally, the standard method of actor training has been to teach an actor how to create gestural modifications appropriate for a character, including changes in posture, facial expression, movement style, and tone of voice (Benedetti, 2012; Brestoff, 1995; Goldstein & Bloom, 2011; Kemp, 2012; Konijn, 2000; Schechner, 2013; Zarrilli, 2009). An actor should take advantage of established gestural codes for producing representations of characters, as based on a character's gender, age, personality, physicality, and so on. An alternative acting method gained popularity in the early part of the 20th century based on the writings of the Russian acting theorist Konstantin Stanislavski (1936), who argued that acting should be predicated not merely on a character's external gestural features, but on his or her internal emotions, beliefs, and motivations (Brestoff, 1995; Cohen, 2004; Goldstein & Bloom, 2011; Hoffman, 2014; Kemp, 2012; Konijn, 2000; Schechner, 2013; Zarrilli, 2009). Stanislavski advocated for a more psychological approach to character portrayal. However, regardless of the method used to get into character, the ultimate objective for the actor is exactly the same: to create a realistic depiction of a person who the actor is not.

The major question that the current study addresses is less about how actors get into a character as about what they modify in their gestural features in order to create realistic depictions that are appropriate for the characters they portray, and how this can be achieved across the wide diversity of characters that any given actor might portray during the course of his or her career. Although it might be relatively straightforward for an actor to play a character who is very similar to him/herself, the real challenge for an actor—and the true art of acting—is to portray characters who are quite different from the self (Smith, 1971). For example, in Shakespeare's day, male actors generally portrayed female characters, and the aim was to hide any indication of their masculinity (Dusinberre, 1998). For a dramatic play like *Romeo and Juliet*, a boy would portray the female character of Juliet. That same actor might portray male characters, such as Puck, in other plays. The issue we want to explore in this study is how actors are able to create portrayals specific for individual characters, and how this operates across multiple distinct characters. In the current experiment, we had professional actors portray a series of nine contrastive characters, including the self.

To examine a wide array of contrastive characters in an experimental study and to do so in a manner that is amenable to quantitative analysis, it is necessary to first have a means of

Matthew Berry and Steven Brown, Department of Psychology, Neuroscience and Behaviour, McMaster University.

This work was supported by a Grant to Steven Brown from the Natural Sciences and Engineering Research Council of Canada.

Correspondence concerning this article should be addressed to Steven Brown, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada. E-mail: stebro@mcmaster.ca

classifying characters. Although the literature of the world is full of stock characters, literary theory provides no systematic classification of them. Indeed, short of generic categories such as round versus flat (Arp & Johnson, 2009) or heroes versus villains (Schmidt, 2001), or likewise intuitive classifications based on a character's gender, age, and species, there has been no classification of characters in literature and drama. To address this shortcoming in the field, as well as to lay the groundwork for analyzing the gestural modifications of actors across contrastive characters, we set out in a previous study to establish a classification scheme for characters. In Berry and Brown (2017), we presented a proposal for a systematic classification of literary characters based on personality dimensions, using a modification of the Thomas-Kilmann Conflict Mode Instrument used in applied studies of personality (Barki & Hartwick, 2001; Duane, 1989; Kilmann & Thomas, 1975, 1977; Rahim, 1985; Thomas, 1992). The Thomas-Kilmann scheme classifies personality along the two orthogonal dimensions of assertiveness and cooperativeness. We conducted a character-rating study in which participants rated 40 stock characters with respect to four personality traits, including assertiveness and cooperativeness. The results showed that ratings of assertiveness and cooperativeness were orthogonal, allowing us to classify characters using these two dimensions. In the current study, we used this classification scheme to organize the series of nine contrastive characters (including the self) with respect to assertiveness and cooperativeness, permitting us to analyze how the prosodic features of an actor's performance vary as a function of these personality dimensions. A major question that we wanted to address was whether a classification of characters based on the subjective assessment of raters (Berry & Brown, 2017) would have behavioral correlates in the prosody of actors portraying these same characters.

Our goal in the present study is to broach the issue of character portrayal by examining the vocal prosody of actors while in character. Only a handful of studies have examined prosodic aspects of character impersonation, although they have primarily focused on simple representations of a particular character or character-type, for example homosexual characters (Cartei & Reby, 2012), Japanese anime characters (Teshigawara, 2004), or fairy tale characters (Doukhan, Rilliard, Rosset, Adda-Decker, & D'Alessandro, 2011). These investigations have only focused on a particular genre of characters and therefore lack a unified perspective on character portrayal as a whole. In addition, the character portrayals in these studies have either come from outside of a lab context using feature films (Cartei & Reby, 2012; Teshigawara, 2004) or from in-lab recordings using nonprofessional actors (Doukhan et al., 2011), and so the present study is, to our knowledge, the first experimental study of character portrayal using professional actors. Relevant research outside of the domain of character portrayal itself includes work on the vocal correlates of lying and of personality traits. For example, studies have shown that lying is reliably associated with an increase in vocal pitch, as well as a greater number of pauses while speaking (Anolli & Ciceri, 1997; Ekman, Friesen, & Scherer, 1976; Villar, Arciulia, & Paterson, 2013). Scherer (1978), in a study of the relationship between personality traits and the perception of voice quality, found that extraversion was associated with loudness. He also showed timbral correlations such that, for example, assertiveness was associated with an absence of breathiness and that agreeableness was associated with an absence of sharpness in the voice.

In contrast to the scant literature on the vocal portrayal of characters, there is a sizable literature on the vocal expression of emotion. This is important since all acting theories agree that a major part of the task of an actor is to convey the emotions of a character (Konijn, 2000). Interestingly, the majority of studies that examine the vocal expression of emotion have used professional actors to portray the emotions (Juslin & Scherer, 2005; Murray & Arnott, 1993). Juslin and Laukka (2003) reported that 87% of a sample of 104 studies examining the vocal expression of emotions employed actors for the creation of stimuli. It is important to point out that using actors to create a stimulus set for a perceptual study of emotion in nonactors is a very different research objective from looking at how a group of actors themselves produce portrayals of these emotions. In a perceptual study, the investigators select that actor's rendition of an emotion that is the most unambiguous and prototypical, whereas in a production study like ours, the use of group-level statistics focuses on the analysis of not only mean trends, but also the variability of depictions across a group of actors. This actor-level variability is missing in perceptual studies.

An analysis of the vocal expression of emotion hinges on how emotions themselves should be classified, but there is at present no single accepted system for representing and classifying emotions. Much research is based on the analysis of so-called basic emotions (Darwin, 1872/1998; Ekman, 1992; Izard, 1992; Plutchik, 2002), such as happiness and sadness (Banse & Scherer, 1996; Goudbeek & Scherer, 2010; Juslin & Laukka, 2003; Juslin & Scherer, 2005; Konijn, 2000; Laukka, 2005; Laukka, Juslin, & Bresin, 2005; Murray & Arnott, 1993; Scherer, 2003; Schröder, 2004). Juslin and Laukka (2003), in a meta-analysis of five broad emotion categories (anger, fear, happiness, sadness, and love-tenderness) across 104 studies of vocal expression, found a number of specific vocal cues that were strongly associated with the five discrete emotions examined. For example, anger and happiness were marked by increases in speech rate, loudness, loudness variability, high-frequency energy (voice quality), mean fundamental frequency (F_0), mean F_0 variability, and (specific to anger) by a small proportion of pauses. In contrast, sadness and tenderness were marked by decreases in speech rate, loudness, loudness variability, high-frequency energy, mean F_0, mean F_0 variability, and (specific to sadness) by a large proportion of pauses.

An alternative system for classifying emotions is to organize them in a dimensional manner. The circumplex model (Russell, 1980, 2003) represents emotions through their placement along the two orthogonal dimensions of arousal (also known as activation; Schröder, 2004; Schröder, Cowie, Douglas-Cowie, Westerdijk, & Gielen, 2001) and valence (also known as evaluation or pleasure; Schröder, 2004; Schröder et al., 2001). There is considerably stronger support for the mapping of prosodic features onto the dimension of arousal than onto that of valence. Previous research has found that arousal is positively correlated with increases in mean F_0 (Goudbeek & Scherer, 2010; Juslin & Scherer, 2005; Laukka et al., 2005; Owren & Bachorowski, 2007; Pereira, 2000; Schröder, 2004; Schröder et al., 2001), F_0 standard deviation/variability (Juslin & Scherer, 2005; Laukka et al., 2005; Schröder et al., 2001), F_0 range (Pereira, 2000; Schröder, 2004; Schröder et al., 2001), mean loudness (Goudbeek & Scherer, 2010; Juslin & Scherer, 2005; Laukka et al., 2005; Pereira, 2000; Schröder, 2004;

Schröder et al., 2001), mean loudness standard deviation/variability (Goudbeek & Scherer, 2010; Laukka et al., 2005), "blaring" timbre (Schröder, 2004), high-frequency energy (Juslin & Scherer, 2005; Schröder, 2004), shimmer (perturbations in amplitude; Goudbeek & Scherer, 2010), speech rate (Juslin & Scherer, 2005; Schröder, 2004), longer phrases (Schröder et al., 2001), and shorter pauses (Juslin & Scherer, 2005; Laukka et al., 2005; Schröder, 2004; Schröder et al., 2001).

In contrast, there is considerably less evidence for the mapping of prosodic features for valence. Parameters correlated with valence are fewer in number, weaker in effect, and more difficult to demonstrate (Bachorowski & Owren, 2008; Owren & Bachorowski, 2007; Pereira, 2000; Schröder, 2004; Schröder et al., 2001). However, the existing evidence suggests that high valence is positively correlated with lower F_0 (Juslin & Scherer, 2005; Laukka et al., 2005; Schröder, 2004), larger F_0 variability (Laukka et al., 2005), larger F_0 range (Juslin & Scherer, 2005; Schröder, 2004), decreased loudness (Juslin & Scherer, 2005; Laukka et al., 2005; Schröder et al., 2001), decreased loudness variability (Goudbeek & Scherer, 2010; Laukka et al., 2005), less high-frequency energy (Schröder, 2004), clearer signal (less timbral noise; Goudbeek & Scherer, 2010), "warm" timbral voice quality (Schröder, 2004), increases in speaking rate (Juslin & Scherer, 2005; Laukka et al., 2005; Schröder, 2004), and shorter pauses (Juslin & Scherer, 2005; Schröder et al., 2001). In general, the dimension of arousal has been found to be more stable than that of valence, with listeners capable of making accurate judgments about a speaker's arousal levels from the same acoustic correlates mentioned above (Douglas-Cowie, Campbell, Cowie, & Roach, 2003; Owren & Bachorowski, 2007; Schröder, 2004).

An important linkage between our dimensional classification of character-types and the dimensional study of emotion is the suggestion that assertiveness is an arousal factor—because it spans from a low-intensity end (unassertive) to a high-intensity end (assertive)—and that cooperativeness is a valence factor, because it spans from a negative end (uncooperative) to a positive end (cooperative). Although this relationship between personality dimensions and emotion dimensions has not be verified experimentally, it opens up the possibility of using the vocal emotion literature to make the predictions that the assertiveness of characters should share prosodic features with the vocal expression of arousal and that the cooperativeness of characters should share features with the vocal expression of valence.

To examine the impact of acting on vocal prosody, we carried out the first experimental study of character portrayal in professional actors, with an emphasis on the modulations of vocal prosody that actors produce to convey specific characters as well as to differentiate characters from one another during performance. (Related work on facial expression will be reported elsewhere.) We took advantage of our previous work (Berry & Brown, 2017) to organize the characters along the two orthogonal personality dimensions of assertiveness and cooperativeness, resulting in a 3×3 scheme based on low, medium, and high values of each dimension (see Figure 1). A major issue that we wanted to address was whether a classification of characters based on the subjective assessment of raters would have behavioral correlates in the vocal prosody of actors portraying these same characters. Professional actors performed the same neutral monologue while portraying each of nine different basic character-types (where the central cell of the scheme was the self). Twelve parameters from the four

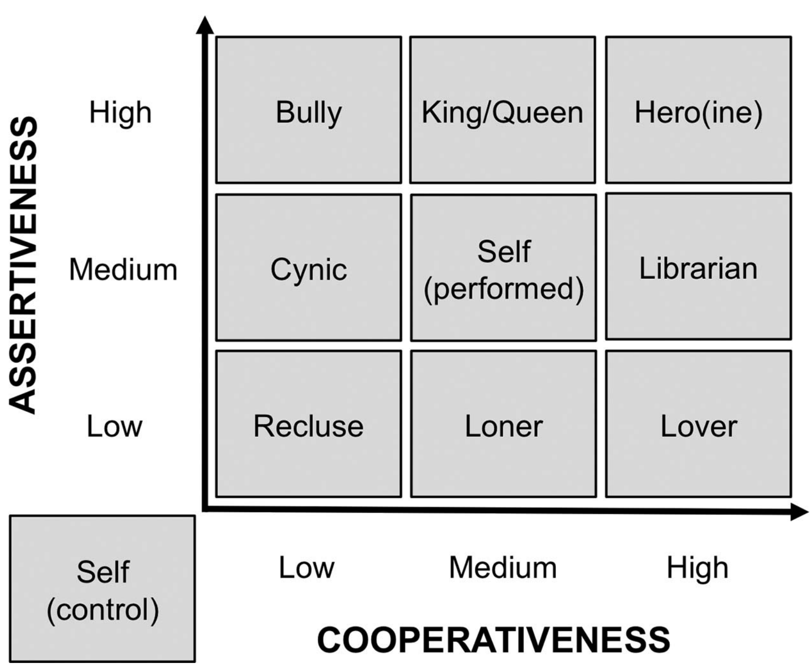

Figure 1. Summary of the nine basic character types portrayed by actors in the study. An orthogonal 3×3 character-classification scheme is presented, as adapted from Berry and Brown (2017). Three levels of assertiveness are crossed with three levels of cooperativeness. The control-self condition at the bottom left is not part of the scheme but instead emerged as a tenth character based on the study's pilot testing.

prosodic categories of pitch, loudness, timbre, and duration/timing were examined based on of the previous literature on vocal prosody (Goudbeek & Scherer, 2010; Juslin & Laukka, 2003; Laukka et al., 2005). Our primary prediction was that the actors' prosodic production would conform to the two-dimensional structure of the character scheme. However, this prediction was tempered by observations in the vocal emotion literature. As a result of the presumed connection between the character dimension of assertiveness and the emotion dimension of arousal, and likewise between the character dimension of cooperativeness and the emotion dimension of valence, we predicted that there would be more and stronger prosodic correlates of assertiveness than of cooperativeness in the vocal data. More specifically, as the level of character assertiveness increases, the actors' portrayals of characters should be higher-pitched, louder, timbrally clearer (less noisy), and rhythmically faster. We also predicted that the prosodic variables would show correlations among themselves, for example the robust relationship between pitch and loudness that is found in the vocal emotion literature (Banse & Scherer, 1996; Goudbeek & Scherer, 2010; Laukka et al., 2005). We examined the results using a combination of univariate methods (analysis of variance) and multivariate methods (principal components analysis) with the overarching goal of identifying the prosodic codes used by actors to vocally embody characters while acting.

Materials and Method

Participants

Twenty-four actors (14 males; 20–63 years; $M = 42.5 \pm 14$ years) were recruited for the experiment through contact with theater companies and academic theater programs in the local and surrounding areas. All actors spoke English either as their native language or fluently as their second language ($n = 1$). They had a mean of 27.5 years ($\pm$ 14.3 years) of acting experience. Fourteen held degrees in acting, with two pursuing degrees in acting at the time of the experiment, and 17 of the 24 self-identified as professional actors. All participants gave written informed consent and were given monetary compensation for their participation. The study was approved by the McMaster University Research Ethics Board.

It is important to point out that the actors were not selected for their training method, but only for their overall level of acting experience. An analysis of an acting-experience questionnaire revealed that about half of the actors had a mixed training that incorporated both gestural and psychological methods, whereas the other half had a more exclusively psychological training (i.e., method acting). As mentioned in the introduction, the current study does not address the issue of how actors get into character, but instead how the prosody of actors varies as a function of the personality dimensions of the portrayed characters. In addition, the single script that was used for all trials (see below) was designed to be emotionally and psychologically neutral, hence providing little to no material for a psychological actor to work with.

Materials

The experiment took place on the stage of a black-box performance laboratory. An 8′ × 10′ foam mat was placed at the center of the stage and was used as the designated performance area. Actors were instructed not to move beyond this area during experimental trials. They performed the tasks facing an empty audience, behind which there was a control room where the experimenter (Matthew Berry) was situated. The experimenter was able to monitor the progress of the session via closed-circuit TV, as well as to provide instructions or answer questions via an intercom. Actors could not see the experimenter while performing. The stage and house lights were adjusted so that the actor was visible on the stage, but that the actor could not see the empty audience seats. The lab's virtual acoustics were adjusted to mimic the acoustics of a performance hall. This was done to increase the ecological validity of the experimental environment.

Actors performed in the designated performance area while being video- and audio-recorded with a Sony XDCam model PXW-X70. The camera was placed 4 m in front of the actor (i.e., in the second row of the audience) on a tripod raised to head level and was zoomed to have the actor's body fill the height of the frame within the designated performance area. Each participant's session was recorded in 1,080 p high-definition video and shot at 60 frames per second. The audio collected from the experimental sessions was extracted for analysis (see the Software and Preprocessing section), while the video recordings were stored for later analysis. As per the recommendations of Titze and Winholtz (1993), which provided suggestions on microphone usage and placement to reduce sampling error and artifacts, video recordings were synchronized with an Apex 575 wireless dual-channel omnidirectional condenser microphone (50–8000 Hz frequency response) attached to an Apex EA01 headset worn by the actor. The microphone transmitter was attached at the waist via a belt, and all wires were covered to reduce restrictions to actor movement. Audio was recorded in stereo at a sampling rate of 48 kHz. The microphone was adjusted to sit parallel to the actor's mouth (i.e., at an angle of around 90 degrees, just beside the cheek) at a distance of 4–10 cm. It moved jointly with the actor, thus eliminating any concerns about measuring vocal loudness. The spectrograms and audio files from each participant's trial were screened for sampling errors and artifacts caused by breathiness, large exhalations, aspirated phonations, or fast head movements. No errors were found.

In addition to recording video, we collected 3D motion-capture recordings of the actors' facial expressions and body gestures. Actors were equipped with 60 infrared reflective markers prior to beginning the experiment (20 on the face, 36 on the torso and limbs, and four on the head via a flexible cap), providing bilateral coverage. Analyses of facial expression and body movement will be reported elsewhere.

Character Scheme

Nine basic characters were performed by the actors, as established by the 3 × 3 (Assertiveness × Cooperativeness) classification scheme validated in Berry and Brown (2017). That rating study showed an absence of stock characters in the bottom-right cell of the scheme (low assertiveness + high cooperativeness). Behavioral piloting of actors' portrayals of stock characters from adjacent cells in the scheme revealed performance differences from the original rating study with regards to the characters of "lover" and "self." They were originally paired with the "librarian" in the medium assertiveness, high cooperativeness cell. However, pilot testing from the current study allowed us to relocate "self" to the central cell and "lover" to the previously unfilled bottom-right cell. The remaining cells in the scheme were filled with basic characters whose locations in the scheme were validated in the rating study. As a result, nine characters (one from each cell, including "lover" and "self") were selected for the present study, as shown in Figure 1.

Actors were given a booklet that listed the names of the nine characters (bully, king/queen, hero/heroine, cynic, self, librarian, recluse, loner, and lover), as well as nine emotions (angry, calm, disgusted, fearful, happy, neutral, proud, sad, and surprised; the emotions are not analyzed here except in one supplementary analysis). There was one character or emotion per page of the booklet, where each character or emotion comprised a single trial. No qualifying information was provided on how to interpret either the characters or the emotions. The order of presentation of the trials was completely randomized across the 18 stimuli for each participant, and no two participants had the same order. Each page consisted of the name of the character (or emotion) that the actor would perform next, in addition to a post-performance questionnaire consisting of four self-rating assessments that the actor filled out immediately after performing a given trial. The four assessments were as follows: (a) rate your satisfaction with your latest performance, (b) rate how deeply you connected to or "felt" the character, (c) rate how well you embodied or portrayed the character, and (d) rate how believable your character portrayal was. All ratings were given on a 7-point Likert scale, spanning from 1 (*not at all*) to 7 (*entirely*). The results of the post-performance assessments will not be discussed in the present paper. Upon completion of the assessment, actors turned the page to initiate the next trial. The session was completed when all 18 trials were performed. The booklet was situated on a music stand at the side of the stage, outside of the designated performance area so as not to restrict movement during performance.

A semantically neutral monologue-script was provided to the actors to memorize in advance of the experiment (see the Appendix). Because of the lack of a scholarly repository of emotionally neutral scripts, we created a custom script for the purposes of this experiment. It was comprised of seven neutral sentences ($M = 6 \pm 1.4$ words/sentence) derived from a set of 10 validated and linguistically neutral sentences (Ben-David, van Lieshout, & Leszcz, 2011). The script was structured such that a small narrative was present (i.e., someone walking into a room and listing the items that they see), but with no indication of emotion or interpersonal interaction. Actors repeated the same monologue for all 18 trials during the session. We decided to use a group of sentences, rather than a single repeated sentence or word (Banse & Scherer, 1996), to better approximate the ecological situation of an acting performance. Debriefing revealed that the actors did not find the script difficult to memorize or perform.

Design and Procedure

The actors were contacted approximately two weeks before their scheduled session and were provided with information pertaining to the study. They were sent the semantically neutral monologue and were instructed to commit it to memory for the experiment. On the day of their session, the actors provided informed consent and completed a participant-information form, acting experience questionnaire, and a brief personality questionnaire. Upon completion of the paperwork, the actors were fitted with motion-capture markers and the wireless microphone and transmitter. The microphone was synchronized and calibrated to the camera. The experimenter then brought the actor to the designated performance area on the stage of the lab and positioned him/her in the center of the foam matting. Actors were instructed to be as expressive and physically active as possible over the course of the session, to move freely anywhere on the matting, to perform to the camera in the audience, and to not turn their back to the camera at any point. They were also instructed to speak clearly and to not whisper.

To acquire a measurement of the participant's normal conversational voice separate from acting trials of the "self" condition (see Figure 1), we used a slight deception. The experimenter, standing beside the video camera in the audience, pretended that the camera required additional calibration and requested that the participant recite the neutral script in a conversational voice as part of the calibration procedure. This recording was used as the "control-self" to compare against the "performed-self" from the acting trial. After the control condition, a practice acting trial (using the character-type of Artist; Berry & Brown, 2017) was then done, after which the experimenter left the theater and did not physically interact with the actor until the session was completed.

The actors were instructed that they could repeat a trial only once if they were dissatisfied with their performance or if they had a false start (i.e., forgot a number of lines). In the case of repeated trials, the second performance was included for analysis. Actors performed the experiment at their own pace and were allowed to take breaks as needed. Each trial lasted roughly 2 min (no time limit was imposed), and the acting phase of the session lasted no more than 45 min. At the end of the session, the actor was debriefed (the deception was explained) and compensated.

Software and Preprocessing

The videos for each session were imported into Adobe Premier Pro CC 2014. Each trial was extracted and exported as a $1{,}920 \times 1{,}080$ .mp4-resolution file. Each trial was 13–131 s in duration ($M = 34.1 \pm 16.4$ s). Once the trial videos had been created, the audio was separated from the video using Abobe Premier Pro CC 2014. Audio samples for each trial were exported as .wav files with 16 bits per sample. These samples were then imported into Praat (Boersma, 2001) for the acoustic analysis of vocal prosody.

A custom Praat script was created to automatically remove all pauses from a given audio sample using a two-pass filter derived from Hirst (2011). The two-pass filter works by initially setting the floor and ceiling of the pitch window liberally to 60 Hz and 700 Hz, respectively, for the first analysis pass. During this pass, the first and third pitch quantiles are extracted from the sample and are used to create a new, more specific pitch window for continued analysis. Specifically, the first quantile is multiplied by a factor of 0.75 to create the new pitch-window floor, and the third quantile is multiplied by a factor of 2 to create the new pitch-window ceiling. The resulting pitch window allows for a more accurate extraction of prosodic data for each speech sample. Pauses were specified as silent moments in the audio sample that had a threshold of 45 dB below the sample mean and a minimum silent interval of 20 ms. A nonpause, or sounding, moment had a minimum sounding interval of 10 ms. Aside from the use of the two-pass filter to specify the pitch window, and the specification of silent/sounding intervals and levels, no other Praat settings were altered for the analyses. The duration parameters were extracted as follows. Pause number was the total number of pauses per trial. To look at timing parameters, the original duration of a trial was broken down into signal duration, corresponding to the total signal or speaking time, and pause duration, corresponding to the total silent time. Using another custom Praat script, we extracted the remaining nine prosodic parameters pertaining to pitch in Hz, loudness in dB, and timbre. Using a two-pass filter again, we extracted the following parameters: the mean, standard deviation, and range of the fundamental frequency in Hz (pitch *M*, pitch *SD*, and pitch range, respectively); the *M*, *SD*, and range of the loudness in dB (loudness *M*, loudness *SD*, loudness range); and the means of the jitter in Hz, shimmer in dB, and noise-to-harmonic ratio (NHR) in dB. The extracted parameters were measured across the entire trial (Banse & Scherer, 1996; Goudbeek & Scherer, 2010).

Data Processing

Normalization of variables. Because of the different measurement units of the 12 dependent variables (see the Software and Preprocessing section), we used a standard z-score normalization procedure to transform and standardize the raw scores at the participant level before the analysis. The formula to calculate a z-score is $z = (x - M)/SD$, where x is the raw score of the participant's variable for a single trial (e.g., the Pitch Mean in Hz for the Bully trial), *M* is the mean score of the participant's variable for all 10 trials (e.g., the averaged pitch *M* across all character trials), and *SD* is the standard deviation of the participant's variable for all 10 trials (e.g., the standard deviation of the averaged Pitch Mean across all character trials). This normalization procedure was applied to each participant across all 10 roles and all 12 parameters. As a result, the data were normalized both per speaker and per parameter.

The per-speaker–per-parameter z-score normalization was chosen over simply normalizing across the parameters alone. Normalizing across each parameter and speaker allows for easier comparison across the data set, as well as for the use of multivariate statistics, such as principal components analysis, because it eliminates participant and gender-related differences at the data level, thereby reducing unwanted participant-induced variance. This transformation also helps satisfy assumptions of normality for the analysis. Similar normalizations have been applied in previous work on emotion and prosody for analyses both within and between studies (Banse & Scherer, 1996; Goudbeek & Scherer, 2010; Laukka, Juslin, & Bresin, 2005).

Univariate analyses. Prior to analysis, each of the 12 normalized parameters was examined for outliers. Any data point beyond a criterion of 2.5 standard deviations away from the mean of a parameter was considered an outlier to be discarded, although no outliers were found for any of the 12 parameters.

Each of the 12 normalized prosodic parameters was analyzed using a linear mixed-effects (LME) regression model with maximum likelihood estimation using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2013). LMEs are similar to repeated-measures analyses of variance in that they control within-subject variance. However, they are not dependent on assumptions or corrections with regards to sphericity and are more robust to missing data. This was useful because two participants whispered during the librarian trials, thus resulting in their data being discarded for these trials. LMEs control within-subject variance through the use of random factors. In the present study, subjects were considered a random factor so as to control for the influence of different subject mean scores on each of the 12 normalized parameters. The two within-subjects factors of assertiveness and cooperativeness were used as fixed effects for the analyses. The control condition (i.e., the control self) was not included in this analysis, whereas the performed self was. The final sample for the univariate analysis was therefore $n = 214$. For the reporting of *F* values, we used a Type III sum of squares with Satterthwaite approximation for degrees of freedom. Statistical significance was set to $\alpha = .05$, and adjustments for repeated testing within a variable group were made using Bonferroni corrections (i.e., for the three parameters within each prosody variable, $\alpha = .0167$; Goudbeek & Scherer, 2010). The significance of statistical analyses and estimates of effect size using eta-squared (η^2), partial eta-squared (η_p^2), and omega squared (ω^2) were calculated using the *afex* package in R (Singmann, Bolker, Westfall, & Aust, 2016).

To compare the two "self" conditions (i.e., performed versus control), we conducted paired sample *t* tests on each of the 12 normalized parameters. Statistical significance was set to $\alpha = .05$ (two-tailed), and adjustments for repeated testing within a variable group were made using Bonferroni corrections (hence $\alpha = .0167$; Goudbeek & Scherer, 2010).

Previous literature on the vocal expression of emotion has shown a strong positive correlation between pitch and loudness (Banse & Scherer, 1996; Belyk & Brown, 2014; Goudbeek & Scherer, 2010); Juslin & Laukka, 2003; Laukka et al., 2005). As a consistency check, we wanted to explore if this relation would hold up using characters, rather than emotions, as the stimulus material. A simple linear regression was calculated to predict the normalized loudness *M* based on the normalized pitch *M* for the nine basic characters and the control. Statistical significance was set to $\alpha =$ .05.

Multivariate analysis. Principal components analysis (PCA) was conducted using the *prcomp* function as a part of the *MVA* (classical MultiVariate Analysis) standard library package in R (R Core Team, 2013). The input data were the 12 normalized parameters, which were zero-centered, but which did not need to be restandardized, because they were all in the same units after normalization. All 10 conditions (i.e., the nine characters plus the control) were included in this analysis for each participant, with the exception of two trials that were omitted due to insufficient data as a result of whispering ($n = 238$; see the Univariate Analyses section). According to Kaiser's criterion (Cangelosi & Goriely, 2007; Dunteman, 1989), the first four PCs should be adequate to describe the majority of variance in a dataset, whereas Cattell's Scree Test and the Broken Stick Method suggest that only the first two PCs are necessary to do this (Cangelosi & Goriely, 2007; Dunteman, 1989). Taking a more conservative approach, we examined the first two PCs. Prior to extraction, and to aid in interpretability, the PCs were rotated using a varimax rotation using the principal function in the *psych* package in R (R Core Team, 2013; Revelle, 2017). The resulting "rotated principal component" (RC) scores are the linear combinations of the 12 normalized parameters, whose weight coefficients are represented by the loadings reported in the Results section.

Results

Univariate Analyses

Main effects of assertiveness and cooperativeness. Table 1 summarizes the results of the linear mixed-effects regression model (LME) analysis for the main effects of character assertiveness and cooperativeness on each of the four prosodic variables and their respective parameters. There was a significant main effect of assertiveness on all 12 dependent variables, suggesting that actors modified their vocal prosody—including the pitch, loudness, timbre, and duration/timing of their speech—across the levels of characters assertiveness. These effects remained significant after Bonferroni corrections for repeated testing.

By contrast, the main effect of cooperativeness on vocal prosody was not nearly as robust. Indeed, only a selection of pitch, loudness, and duration/timing parameters showed significant main effects of cooperativeness, as timbre parameters failed to reach significance. Bonferroni correction eliminated the effect on pitch, $F(2, 214) = 3.26$, $p = .029$, leaving only weak effects on loudness and duration/timing (see Table 1). This finding is supported by the significantly lower estimates of effect size for cooperativeness when compared to assertiveness. Overall, prosodic changes were much more strongly associated with the assertiveness level of a character than its cooperativeness level.

Table 1
Linear Mixed-Effects Regression Model Analysis of Variance

Prosody and parameter	Source of variation	*df*	Type III SS	Mean square	*F* value	*p*-value	η^2	η_p^2	ω^2
Pitch									
Pitch *M*	Assertiveness	2	91.18	45.59	101.09	<.001***	.47	.49	.47
	Cooperativeness	2	3.26	1.63	3.61	.029*	.02	.03	.01
	Assert. × Coop.	4	1.79	.45	.99	.412	.01	.02	.00
Pitch *SD*	Assertiveness	2	84.54	42.27	84.85	<.001***	.43	.44	.43
	Cooperativeness	2	1.75	.88	1.76	.175	.01	.02	.00
	Assert. × Coop.	4	2.90	.72	1.45	.217	.01	.03	.00
Pitch range	Assertiveness	2	62.03	31.01	52.29	<.001***	.32	.33	.31
	Cooperativeness	2	.28	.14	.24	.788	.00	.00	.00
	Assert. × Coop.	4	4.98	1.25	2.10	.082	.03	.04	.01
Loudness									
Loudness *M*	Assertiveness	2	122.19	61.09	191.59	<.001***	.61	.64	.60
	Cooperativeness	2	3.37	1.69	5.29	.006**	.02	.04	.01
	Assert. × Coop.	4	8.22	2.06	6.45	<.001***	.04	.11	.03
Pitch *SD*	Assertiveness	2	21.55	10.78	14.26	<.001***	.11	.12	.10
	Cooperativeness	2	1.07	.53	.71	.494	.01	.01	.00
	Assert. × Coop.	4	6.60	1.65	2.18	.072	.03	.04	.02
Loudness range	Assertiveness	2	103.00	51.50	126.85	<.001***	.52	.54	.52
	Cooperativeness	2	3.99	1.99	4.91	.008**	.02	.04	.02
	Assert. × Coop.	4	3.27	.82	2.02	.093	.02	.04	.01
Timbre									
Jitter	Assertiveness	2	25.12	12.56	21.13	<.001***	.14	.16	.14
	Cooperativeness	2	3.25	1.63	2.73	.067	.02	.02	.01
	Assert. × Coop.	4	19.05	4.76	8.01	<.001***	.11	.13	.10
Shimmer	Assertiveness	2	32.05	16.02	27.58	<.001***	.18	.20	.17
	Cooperativeness	2	.49	.24	.42	.658	.00	.00	.00
	Assert. × Coop.	4	22.13	5.53	9.52	<.001***	.12	.15	.11
NHR	Assertiveness	2	28.68	14.34	21.82	<.001***	.16	.17	.15
	Cooperativeness	2	3.02	1.51	2.30	.103	.02	.02	.01
	Assert. × Coop.	4	8.26	2.07	3.14	.015*	.05	.06	.03
Duration/timing									
Signal duration	Assertiveness	2	21.37	10.69	15.21	<.001***	.12	.13	.11
	Cooperativeness	2	7.82	3.91	5.56	.004**	.04	.05	.03
	Assert. × Coop.	4	9.21	2.30	3.28	.012*	.05	.06	.03
Pause duration	Assertiveness	2	35.75	17.88	34.46	<.001***	.22	.25	.21
	Cooperativeness	2	3.97	1.98	3.83	.023*	.03	.04	.02
	Assert. × Coop.	4	13.88	3.47	6.69	<.001***	.08	.11	.07
Pause number	Assertiveness	2	38.87	19.44	33.18	<.001***	.22	.24	.22
	Cooperativeness	2	.63	.32	.54	.585	.00	.01	.00
	Assert. × Coop.	4	10.02	2.51	4.28	.002**	.06	.07	.04

Note. Assert. = assertiveness; Coop = cooperativeness; NHR = noise-to-harmonic ratio. Summary of within-subject effects of assertiveness and cooperativeness on the 12 dependent variables. Analysis of variance table with Type III sum of squares (SS) using Satterthwaite approximation for degrees of freedom ($n = 216$). Denominator $df = 214$. Measures of effect size include η^2, eta-squared; η_p^2, partial eta-squared, and ω^2, omega squared.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 2 plots the means of each acoustic parameter in order to further investigate the differences between the levels of assertiveness. The main effect of assertiveness was characterized by monotonic increases in both the pitch and loudness variables, with the six dependent variables demonstrating a low < medium < high increasing trend. As the level of assertiveness of characters increased, vocal expressions became higher pitched and louder, with more variability and range in these parameters. Different effects were observed for timbre and duration/timing. As the level of assertiveness increased, both timbre and duration parameters decreased, with timbre parameters decreasing in a low > medium > high trend (i.e., the character's voices became clearer due to fewer frequency and amplitude perturbations), and duration decreasing in a low > medium = high trend (i.e., characters spoke faster with fewer/shorter pauses). This effect demonstrates that, as the assertiveness level of a character increased, the vocal expressions became clearer and the utterances shorter.

Figure 3 shows the same analysis for cooperativeness. The LME analysis revealed that the main effect of cooperativeness after correcting for repeated testing was characterized by loudness and duration/timing parameters, specifically loudness *M*, loudness range, and signal duration. The figure demonstrates a slightly decreasing trend in a low > medium > high fashion with regards to loudness *M*, as well as a decreasing trend in a low > medium = high fashion with regards to loudness range. These effects show that, as the cooperativeness level of character increased, the actor's mean loudness and loudness range decreased. The pattern was different for Signal Duration. As the cooperativeness level of characters increased, the signal duration decreased and then increased in a medium < low < high fashion. This nonmonotonic effect may reflect idiosyncratic features of the specific characters chosen for the study.

Interactions and simple effects of assertiveness and cooperativeness. The LME analysis revealed several interaction effects between assertiveness and cooperativeness across the prosodic variables of loudness, timbre, and duration, even after correcting for repeated testing. There were no significant interaction effects for pitch. The parameters that did show interactions were

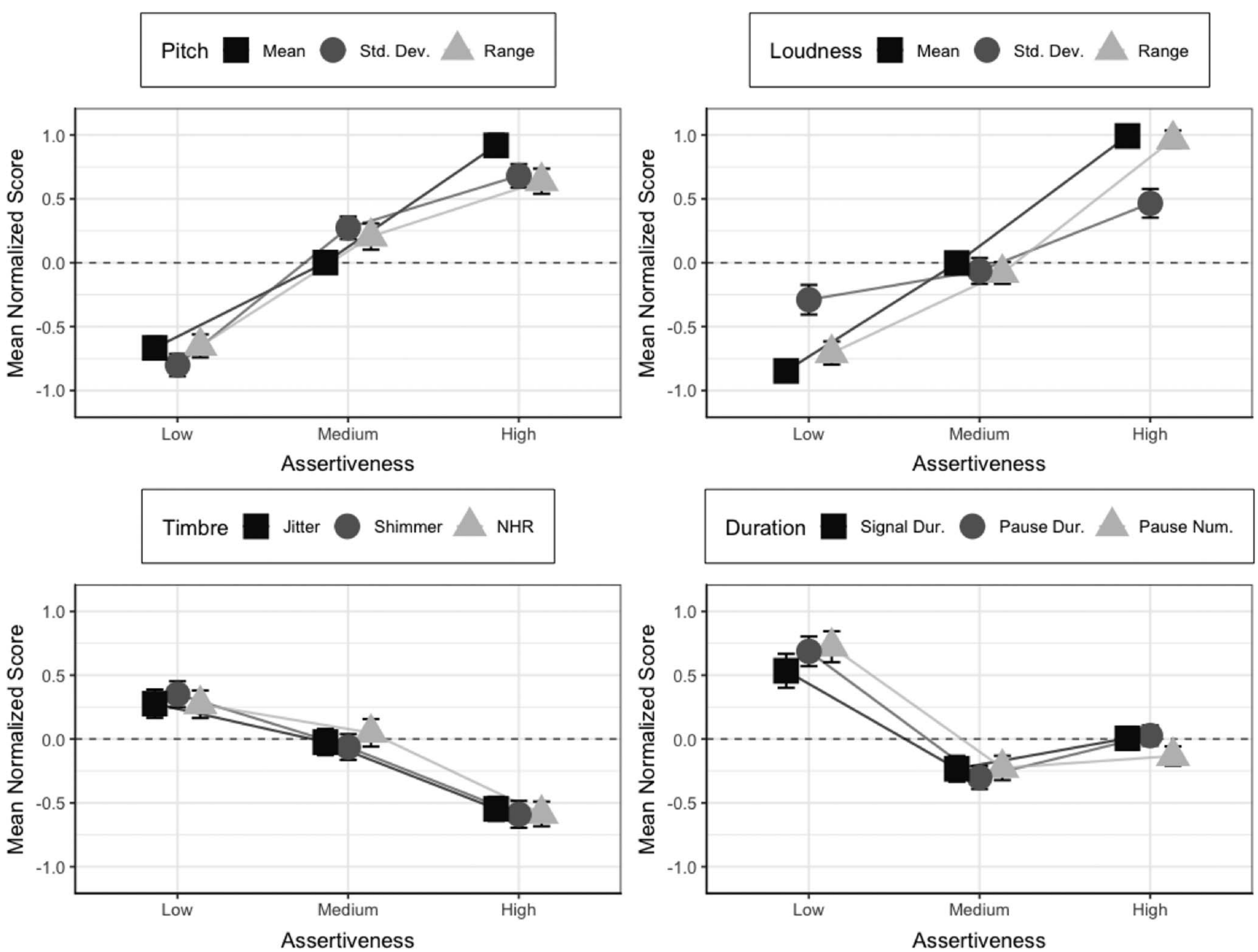

Figure 2. Effect of character assertiveness on the 12 prosodic parameters. Axis values are calculated by first averaging the subject-normalized scores across character and then across level of assertiveness. Dur. = duration; NHR = noise-to-harmonic ratio; Num = number (of pauses); Std. Dev. = standard deviation. Error bars show the standard error of the mean.

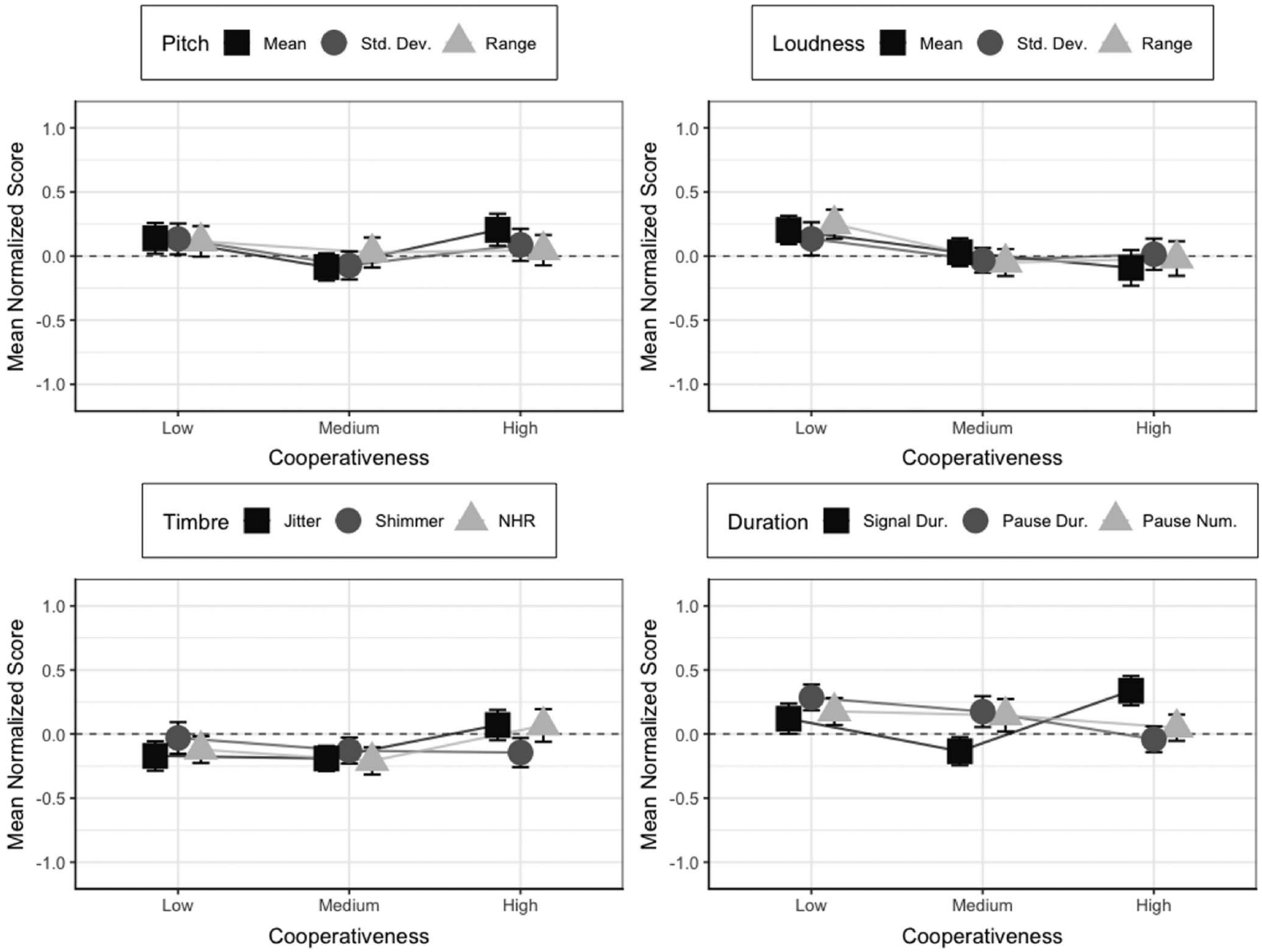

Figure 3. Effects of character cooperativeness on the 12 prosodic parameters. Axis values are calculated by first averaging the subject-normalized scores across character and then across level of cooperativeness. Dur. = duration; NHR = noise-to-harmonic ratio; Num = number (of pauses); Std. Dev. = standard deviation. Error bars show the standard error of the mean.

loudness *M*, $F(4, 214) = 6.5$, $p < .001$; jitter, $F(4, 214) = 8.0$, $p < .001$; shimmer, $F(4, 214) = 9.5$, $p < .001$; NHR, $F(4, 214) = 3.4$, $p = .015$; signal duration, $F(4, 214) = 3.3$, $p = .023$; pause duration, $F(4, 214) = 6.7$, $p < .001$; and pause number, $F(4, 214) = 4.3$, $p = .002$. The interaction effects and effect sizes are summarized in Table 1.

To further investigate the interaction effects of the seven parameters, the data were split and averaged by the three levels of cooperativeness, thereby allowing the pairwise comparison of each level of assertiveness within each level of cooperativeness, correcting for multiple corrections. These simple effects are statistically summarized in Supplementary Table S1 in the online supplementary material, and the text after the table describes specific interactions in detail. Only the effects for loudness *M* will be reported here. The simple effects of assertiveness for loudness *M* revealed a L (low) < M(medium) = H (high) pattern at low cooperativeness, a L < M < H pattern at medium cooperativeness, and a L = M < H pattern at high cooperativeness, indicating a decreasing shift in loudness across performances, and making the librarian character similar to the lover with regards to mean loudness across the performance. Due to the structure of the predicative scheme, an interaction effect and simple effects analysis between basic character-types results in a direct comparison between the two basic characters within a given category (in this case it will be cooperativeness). The results thus provide a more specific prosodic profile of the average character type or types being examined.

Analysis of performed-self versus control-self. An incidental finding of the pilot testing of the study was that the "self" condition appeared to be performed as a character, with a different prosody than an actor's conversational voice. It is for this reason that we introduced the microphone-calibration deception described in the Design and Procedure section to obtain the participant's conversational prosody while reciting the script. Table 2 summarizes the results of paired-sample *t* tests comparing the performed-self condition to the control-self condition. There were significant differences between conditions for all dependent variables, with

Table 2
Paired-Sample T-Tests for Performed-Self Versus Control-Self

Prosody and parameter	Condition	*M*	*SD*	*t*-value	*p*-value	Cohen's *d*
Pitch						
Pitch *M*	Self (control)	−.76	.50	5.75	<.001***	1.17
	Self (performed)	−.06	.49			
Pitch *SD*	Self (control)	−.44	.76	4.37	<.001***	.89
	Self (performed)	.20	.59			
Pitch range	Self (control)	−.56	.70	4.31	<.001***	.88
	Self (performed)	.36	.85			
Loudness						
Loudness *M*	Self (control)	−.43	.57	4.74	<.001***	.97
	Self (performed)	.13	.43			
Loudness *SD*	Self (control)	−.34	.93	.93	.361	.19
	Self (performed)	−.11	.66			
Loudness range	Self (control)	−.54	.64	2.32	.029*	.47
	Self (performed)	−.09	.61			
Timbre						
Jitter	Self (control)	.88	.91	−3.31	.001***	.68
	Self (performed)	.09	.61			
Shimmer	Self (control)	.90	.76	−4.04	.001***	.83
	Self (performed)	.20	.47			
NHR	Self (control)	.80	.85	−3.05	.006**	.62
	Self (performed)	.13	.71			
Duration/timing						
Signal duration	Self (control)	−.94	.23	1.34	.194	.27
	Self (performed)	−.74	.67			
Pause duration	Self (control)	−1.26	.53	3.16	.004**	.64
	Self (performed)	−.75	.50			
Pause number	Self (control)	−1.11	.50	2.53	.019*	.52
	Self (performed)	−.61	.71			

Note. NHR = noise-to-harmonic ratio. Degrees of freedom = 23 for all tests. For Cohen's *d*, small effect = .2, medium effect = .5, large effect = .8. * $p < .05$. ** $p < .01$. *** $p < .001$.

the exception of loudness *SD* and signal duration. After applying Bonferroni correction, loudness range, $t(23) = 2.32$, $p = .029$, and pause number, $t(23) = 2.53$, $p = .019$, became nonsignificant. These results indicate that an actor's voice was significantly different when doing a stage performance of him- or herself compared to when talking in his or her conversational voice. Specifically, performing the role of oneself on stage resulted in a prosody that was higher-pitched, louder, slower, and timbrally clearer than one's conversational voice.

Correlation between pitch and loudness. A simple linear regression was calculated to look at the relation between normalized pitch *M* and normalized loudness *M* (see Figure 4). A significant regression equation was found, $F(1, 236) = 310.6$, $p < .001$, with an r value of .75, and an adjusted R^2 value of .57. This is consistent with previous findings in the study of vocal emotion showing that pitch and loudness are strongly positively correlated (Banse & Scherer, 1996; Goudbeek & Scherer, 2010). This creates an important prosodic parallel between character portrayal and the expression of emotion. This result remained consistent across the different character-types. In splitting the data by character-type, significant positive correlations persisted between normalized pitch *M* and normalized loudness *M* for all but two characters, the bully and the loner (see Supplementary Table S2 and Supplementary Figure S1 in the online supplementary material).

Multivariate Analyses

Correlations. Prior to running the PCA, a Pearson product–moment correlation analysis was conducted to determine if any prosodic parameters were too weakly or strongly correlated to be used in the multivariate analysis so as to avoid any block diagonal correlation matrices (Dunteman, 1989). The results of the correlation analysis showed that no variables needed to be excluded from the PCA (see Table 3).

A number of parameters were significantly correlated with one another. In addition to high levels of correlation within a given variable (e.g., pitch parameters being highly positively correlated with other pitch parameters), there was a range of weak to strong correlations between variables as well. For example, pitch *M* was significantly positively correlated not only with other pitch parameters (pitch *SD*, $r = .80$, $p < .001$; pitch range, $r = .80$, $p < .001$), but with loudness parameters (loudness *M*, $r = .75$, $p < .001$; loudness *SD*, $r = .37$, $p < .001$; loudness range, $r = .74$, $p < .001$). It was significantly negatively correlated with timbral parameters (jitter, $r = -.39$, $p < .001$; shimmer, $r = -.51$, $p < .001$; NHR, $r = -.48$, $p < .001$) and durational parameters (pause duration, $r = -.14$, $p < .05$; pause number, $r = -.15$, $p < .05$). Furthermore, although the pitch and loudness variables were positively correlated with one another, both timbral and duration variables were negatively correlated with all other variables.

PCA. Table 4 presents a summary of the PCA results. The first two PCs were extracted using a varimax rotation. The bottom part of Table 4 describes how the resulting RCs accounted for 61.4% of the total variance in the dataset, with RC1 accounting for around 42.2% of the total and RC2 around 19.2% of the total.

The interpretation of the RCs can be taken from the respective loadings, as summarized in the upper part of Table 4. RC1 had

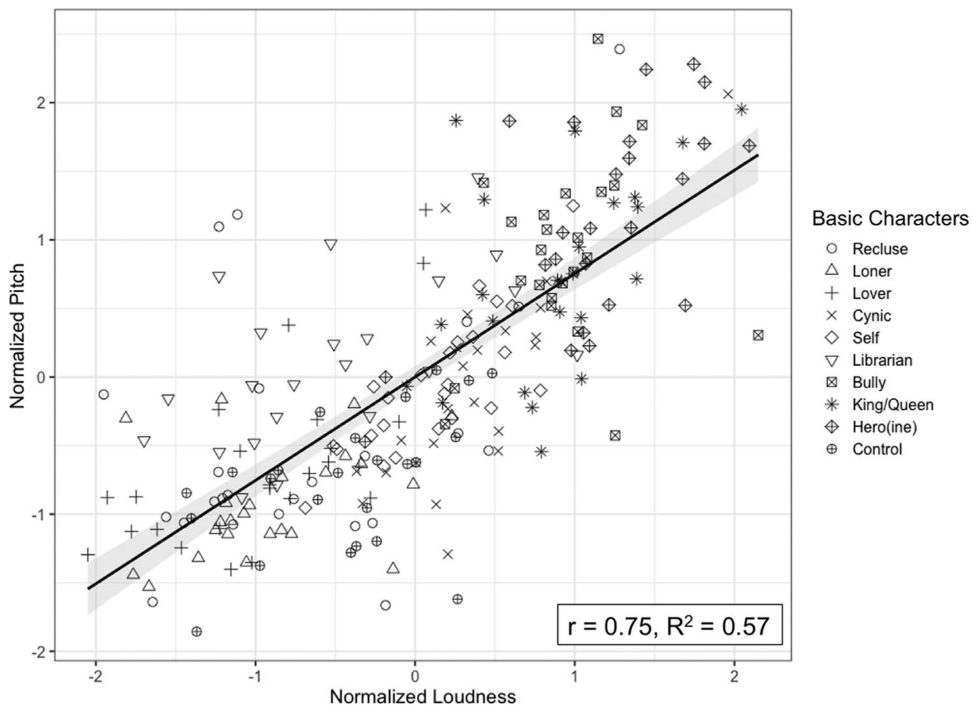

Figure 4. Linear regression and correlation of normalized pitch and loudness. Pitch mean and loudness mean values were normalized within participants by applying a z-score transformation. This included data from the control-self trials. $F(1, 236) = 310.6$, $p < .001$, $r = .754$, $R^2 = .568$. Each point represents an observation from a single trial. Characters are represented by different shapes. The regression line is depicted in black, whereas the shaded area indicates the standard error of the regression line.

high loadings for all of the pitch, loudness, and timbral parameters, whereas RC2 had high loadings for all of the duration/timing parameters. These loadings are consistent with the intervariable correlations reported in the previous section. Figures 5 and 6 plot the first two RCs and their generated component scores. They are color-coded to show the three levels of assertiveness and cooperativeness, respectively. The first RC appears to account for variation in the levels of assertiveness, while the second RC appears to account for variation in performance (i.e., performance of characters versus the nonperformed control). There does not seem to be

Table 3
Correlation Matrix for the 12 Prosodic Parameters

Prosody and parameter	f0M	f0SD	f0R	LdM	LdSD	LdR	Jit	Shim	NHR	SDur	PDur
Pitch											
f0M	—										
f0SD	$.80^{***}$	—									
f0R	$.73^{***}$	$.88^{***}$	—								
Loudness											
LdM	$.75^{***}$	$.64^{***}$	$.58^{***}$	—							
LdSD	$.37^{***}$	$.29^{***}$	$.28^{***}$	$.24^{***}$	—						
LdR	$.74^{***}$	$.62^{***}$	$.56^{***}$	$.84^{***}$	$.43^{***}$	—					
Timbre											
Jit	$-.39^{***}$	$-.18^{**}$	$-.16^{*}$	$-.55^{***}$	$-.20^{**}$	$-.53^{***}$	—				
Shim	$-.51^{***}$	$-.38^{***}$	$-.34^{***}$	$-.55^{***}$	$-.26^{***}$	$-.50^{***}$	$.71^{***}$	—			
NHR	$-.48^{***}$	$-.25^{***}$	$-.22^{***}$	$-.50^{***}$	$-.25^{***}$	$-.47^{***}$	$.83^{***}$	$.76^{***}$	—		
Duration											
SDur	.04	−.02	.05	$-.23^{***}$	$.28^{***}$	−.04	.02	$-.14^{*}$	−.01	—	
PDur	$-.14^{*}$	$-.18^{**}$	$-.16^{*}$	$-.17^{*}$	$-.15^{*}$	$-.14^{*}$	$-.15^{*}$	−.08	$-.14^{*}$	$.28^{***}$	—
PNum	$-.15^{*}$	$-.14^{*}$	−.10	$-.33^{***}$	.04	$-.20^{**}$	.05	−.02	−.01	$.56^{***}$	$.68^{***}$

Note. f0M = pitch *M*; f0SD = pitch *SD*; f0R = Pitch Range; LdM = loudness *M*; LdSD = loudness *SD*; LdR = loudness range; Jit = jitter; Shim = shimmer; NHR = noise-to-harmonic ratio; SDur = signal duration; PDur = pause duration; PNum = number of pauses. $n = 238$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4
Varimax-Rotated Principal Component Loading Matrix

Prosody and parameter	Rotated components RC1	RC2	Communalities	Uniqueness
Pitch				
Pitch *M*	**.86**	−.18	.78	.22
Pitch *SD*	**.74**	−.31	.65	.35
Pitch range	**.69**	−.28	.56	.44
Loudness				
Loudness *M*	**.85**	−.29	.80	.20
Loudness *SD*	**.46**	.07	.22	.78
Loudness range	**.85**	−.17	.75	.25
Timbre				
Jitter	**−.70**	−.30	.58	.42
Shimmer	**−.77**	−.29	.67	.33
NHR	**−.73**	−.33	.64	.36
Duration/timing				
Signal duration	.05	**.63**	.40	.60
Pause duration	−.05	**.79**	.62	.38
Pause number	−.11	**.82**	.69	.31
Eigenvalue	5.07	2.30		
% of explained variance	42.20	19.20		
% of cumulative variance	42.20	61.40		

Note. NHR = noise-to-harmonic ratio. Numbers in bold indicate significant loadings of the RC's by the prosodic parameter listed.

a relation between either RC and levels of cooperativeness. Therefore, the first two RCs can be interpreted as reflecting, respectively, a dynamic vocal factor (RC1) and a performative factor (RC2). In general, it appears that actors modulate pitch, loudness, and timbral parameters in a combinatorial manner to differentiate between levels of character assertiveness (with higher, louder, and clearer values indicative of higher assertiveness) and that they modulate duration/timing parameters to differentiate between performed speech and regular speech (with slower utterances indicative of the former).

The character/emotion relationship. Finally, as a purely exploratory analysis, we investigated the relationship between character dimensions and emotion dimensions in order to see if the connection between assertiveness and arousal, on the one hand, and cooperativeness and valence, on the other, was obtained. Supplementary Figure S2 in the online supplementary material shows a PCA plot of the 10 characters (including the control-self) and the nine emotions tested. High-assertive characters like the king, hero, and bully clustered with high-arousal emotions like happy, angry, and surprised. The control-self clustered with neutral emotion. And low-arousal characters like the loner, recluse, and lover clustered with low-arousal emotions like calm and sad, although they were also proximate to higher-intensity emotions such as fear and disgust. The association between the personality dimension of assertiveness and the emotion dimension of arousal may be underlain by physiological mechanisms related to the level of activation of the autonomic nervous system (Banse & Scherer, 1996; Goudbeek & Scherer, 2010).

Discussion

In this first experimental study of character portrayal using professional actors, we found that actors modulated the pitch, loudness, timbre, and timing properties of their vocal productions in order to conform with the personality traits of characters. The effect was much more robust for assertiveness—which has similarities to the arousal dimension of emotion—than for cooperativeness, which has similarities to the valence dimension of emotion. Hence, this mirrors a general finding from the vocal-emotion literature that arousal is much more reliably encoded than is valence. Next, we reported the incidental finding that an analysis of the performed-self showed important prosodic differences from an individual's conversational voice. The performed-self was shown to be higher-pitched, louder, and slower than one's conversational manner of speaking, reflecting what we will refer to below as a "performance persona" of the self. Finally, looking beyond the character scheme, we found that key prosodic variables were correlated with one another across characters, including pitch and loudness, again paralleling findings from the vocal-emotion literature. Overall, the results show that actors employ prosodic codes in order to embody specific characters based on their personality dimensions. The results also show that a classification of characters based on the subjective assessment of raters (Berry & Brown, 2017) had clear behavioral correlates in the prosody of actors portraying these same characters.

The Personality of Characters Predicts Actors' Vocal Prosody

According to the view of acting as mimesis first espoused by Plato (380 BCE/1968), actors engage in processes of personal mimicry in order to recreate prototypical properties of characters based on their presumed personality and physical features (Kemp, 2012; Schechner, 2013; Zarrilli, 2009). In the present study, we established predictions for some of these features by organizing the selected characters according to a two-dimensional personality scheme that had been previously validated in a study from the lab (Berry & Brown, 2017). By doing so, we were able to provide

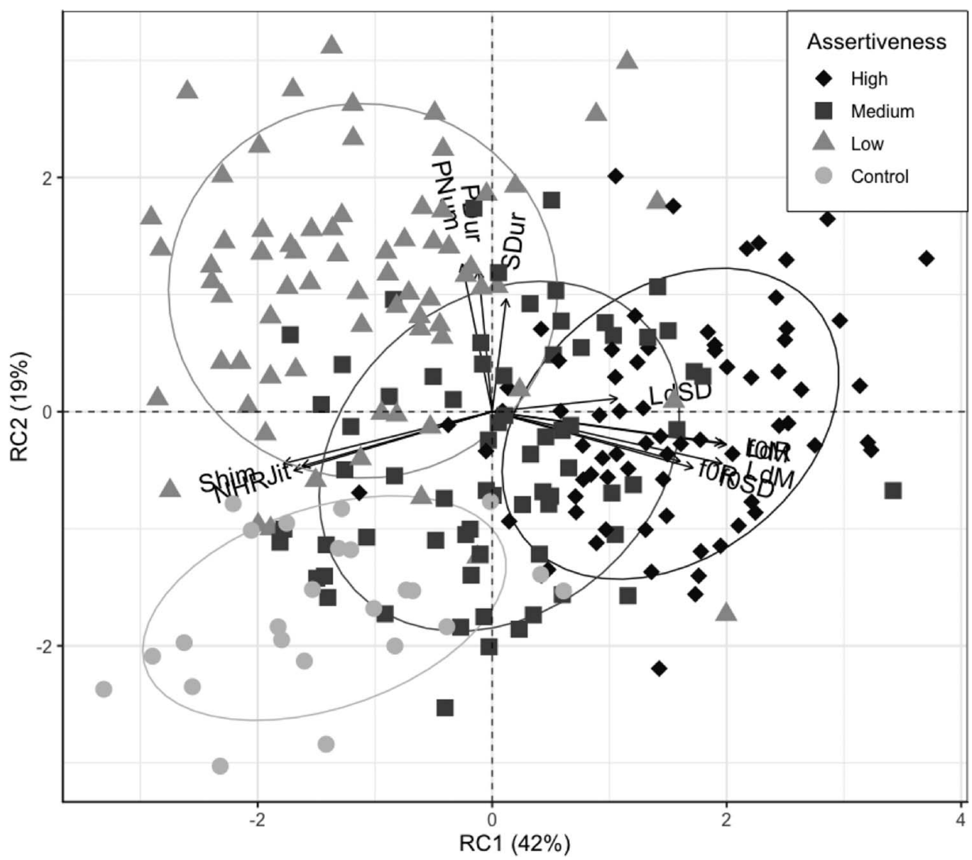

Figure 5. Varimax rotated principal components biplot for levels of assertiveness. Biplot axes of the first two extracted varimax rotated principal components account for 42.2% and 19.2% of the explained variance, respectively. Ellipse probability for levels of assertiveness is 70%. f0M = pitch *M*; f0SD = pitch *SD*; f0R = pitch range; LdM = loudness *M*; LdSD = loudness *SD*; LdR = loudness range; Jit = jitter; Shim = shimmer; NHR = noise-to-harmonic ratio; SDur = signal duration; PDur = pause duration; PNum = number of pauses.

support for the notion that assertiveness is far more predictive of actors' vocal prosody than is cooperativeness, a finding that matches the perceptual advantage for recognizing arousal more reliably than valence in studies of emotion discrimination (Belyk & Brown, 2014; Douglas-Cowie et al., 2003; Goudbeek & Scherer, 2010; Owren & Bachorowski, 2007; Schröder, 2004). This was demonstrated both as main effects in the univariate analyses and by the separation of low, medium, and high levels of assertiveness in the PCA analysis, while failing to find such a separation for the levels of cooperativeness. The prosodic correlates of cooperativeness were few in number and weak in effect. Only three parameters remained significant after correcting for multiple comparisons. Increases in cooperativeness were associated with decreases in loudness and loudness range. Signal duration displayed a nonmonotonic relation to cooperativeness.

Assertiveness, by contrast, showed effects across the spectrum of prosodic parameters. Increases in character assertiveness were associated with increases in pitch, loudness, voice clarity, and speech rate. These findings correspond with arousal-related effects in the vocal-emotion literature, perhaps with the exception of timbre (Goudbeek & Scherer, 2010; Juslin & Scherer, 2005; Schröder, 2004). However, the current understanding of timbral effects on arousal-related vocal-expression is limited, with some evidence suggesting weak, indirect relationships (Gobl & Ní Chasaide, 2003; Goudbeek & Scherer, 2010; Owren & Bachorowski, 2007). Hence, there seems to be a general concordance between our findings about character portrayal and those about vocally expressed emotions. The interpretation of such a relationship is not that characters can be equated with specific emotions (i.e., flat characters; Forster, 1927/2005), but that the personality of a character may bias him/her to experience or display some emotions more than others. For example, the hero might show a tendency to experience high-arousal and positive-valenced emotions, while the bully might show a tendency to experience high-arousal and negative-valenced emotions. However, at another level, characters and emotions might interact in a more combinatorial manner. For example, there are happy kings and sad kings, and even a congenitally happy king can become sad when a loved one dies. Particular character-archetypes can thus be manifested in characters showing a diversity of emotional traits, and even a character who predominantly shows one particular emotional trait throughout a story can experience a variety of emotional states, including those that defy their general personality. Therefore, the relationship between characters and emotions should be thought of as a complex one-to-

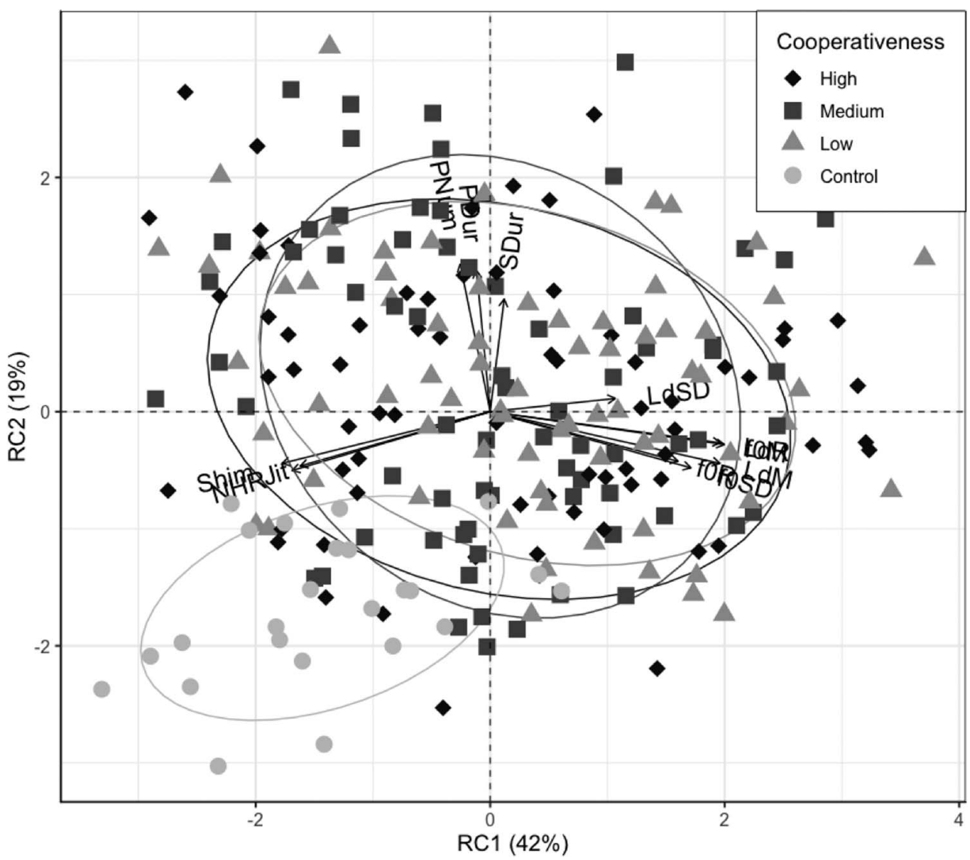

Figure 6. Varimax rotated principal components biplot for levels of cooperativeness. Biplot axes of the first two extracted varimax rotated principal components account for 42.2% and 19.2% of the explained variance, respectively. Ellipse probability for levels of cooperativeness is 70%. f0M = pitch *M*; f0SD = pitch *SD*; f0R = pitch range; LdM = loudness *M*; LdSD = loudness *SD*; LdR = loudness range; Jit = jitter; Shim = shimmer; NHR = noise-to-harmonic ratio; SDur = signal duration; PDur = pause duration; PNum = number of pauses.

many relationship, not a one-to-one relationship. Characters that are archetypal will tend to have more-limited emotional profiles than characters that are more complex. It is probably for this reason that our study of archetypal characters is showing a strong dimensional relationship with emotions. This relationship might turn out to be weaker in a study using more-complex characters.

The Performed-Self

An incidental finding of the present study was that the self that was performed during the acting trials had different prosodic parameters than those of the conversational self during control trials. This was shown in both the univariate analyses and the PCA analyses, where the second principal component effectively distinguished the control-self from all of the acted conditions. Despite the fact that the actors were provided with the opportunity to drop their performance façade and to speak the script conversationally, they remained "on" during the self trials and recited the script as a more exaggerated version of themselves. We are calling this effect the adoption of a "performance persona", in accordance with dramaturgical perspectives in social psychology (Goffman, 1959; Landy, 1993; Shulman, 2017), which argue that the self is comprised of a number of context-specific variants that can be considered as personas of the self during everyday social interactions. To the best of our knowledge, this is the first reported instance of a direct comparison between self conditions with regards to prosodic parameters. Acoustically, the performed-self was shown to be higher pitched, louder, clearer, and slower than the conversational manner of speaking. In this regard, it showed strong similarities with the main effect of assertiveness, with the exception of the timing parameters, where the performed-self was intoned more slowly than the control-self, with longer pause duration and a greater number of pauses. Hence, the performed-self is an assertive persona in which speech rate is reduced, perhaps as a form of demonstration.

The performed-self shows tantalizing similarities with other well-known prosodic phenomena in the literature, most especially infant-directed speech (sometimes called "motherese"), in other words the manner in which caregivers interact vocally with their babies. Infant-directed speech is characterized by the same prosodic suite of higher pitch, larger pitch-variability, increased loudness, and slower tempo as compared to adult-directed speech (Fernald, 1989; Papoušek, Papoušek, & Symmes, 1991; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011; McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013). We believe that this combination of features comprises a performance persona, which is expressed when a solo speaker is interacting with an attentive audience that is generally listening passively, rather than engaging in a dialogue. Such is the characteristic situation of caregiver–infant interaction but is also the discursive arrangement of a seminar speaker, a tour guide, the narrator of a story, and many other situations where one speaker plays a dominant role in an interaction with attentive, but typically silent, recipients. The prosody of the performance persona is designed not only to capture attention (high pitch, high amplitude; Brosch, Grandjean, Sander, & Scherer, 2008), but also to be demonstrational (reducing ambiguity via slow tempo and the use of pausing; Snedeker & Trueswell, 2003). In the case of infant-directed speech, it also supports language learning (Ma et al., 2011) and social bonding (Benders, 2013). Overall, this is a type of prosody that is optimized for situations where one person is actively conveying important information to one or more recipients of this information. It is this communicative goal that most likely unites the prosodies of a caregiver and a university lecturer.

Correlations Among Prosodic Variables

In addition to looking at the predictive potential of the character scheme, we wanted to examine correlations among the prosodic parameters themselves. Consistent with findings from the vocal-emotion literature (Banse & Scherer, 1996; Goudbeek & Scherer, 2010; Scherer, 2003), we found a strong and significant correlation between pitch and loudness across character trials. This correlation is known in the performance literature as the “high-loud” rule of expression (Friberg, Bresin, & Sundberg, 2006), indicating that an increase in pitch is often accompanied by an increase in loudness (Belyk & Brown, 2014; Goudbeek & Scherer, 2010; Laukka et al., 2005), both vocally and instrumentally (Friberg et al., 2006). Our results indicate that this expression rule is very robust, being active not only in the vocal expression of emotions, but also in the portrayal of characters during acting. Hence, it might reflect a general physiological relationship between pitch and loudness during sound production.

Limitations

There are a number of limitations of this study. First, only a select number of basic characters could be examined within the scope of this performance study. Berry and Brown (2017) established a corpus of 40 basic characters to draw upon for the present study, of which only nine were used. However, this is comparable to the number of emotions that are tested in many studies of vocal emotion (Juslin & Scherer, 2005; Laukka, 2005; Murray & Arnott, 1993; Schröder, 2004). Next, because of the exploratory nature of the study, we wanted to start out with a set of simple, rather than complex, characters. Hence, all of the characters were basic, archetypal character types, rather than complex and/or more realistic characters like Romeo or Juliet. Despite this, our use of a dimensional approach ensures that future acting studies using more-complex characters can generate predictions about character prosody based on features like personality traits. This can be useful to account for (a) the expression of different personality traits for a single character over the course of a drama, and/or (b) relative differences among the dramatis personae of a story. In *Romeo and Juliet*, for example, Romeo goes from being a prototypical lover to being a vengeful hero when he kills Tybalt in response to Tybalt’s slaying of his best friend. Likewise, Romeo and Tybalt have starkly different personalities between themselves throughout the drama; they represent different character archetypes. These features of drama and literature can be better addressed using a dimensional approach than by using one that views all characters as singular, unrelated entities.

Another limitation of the study relates to the categories of traits that were used for the predictive scheme. Although they turned out to be effective at predicting prosodic changes in the current analysis (mainly for assertiveness), the personality dimensions of assertiveness and cooperativeness are by no means the only personality traits that are relevant for describing characters. Other important traits include intelligence, extraversion, introversion, to name a few, or even valence and arousal more directly. Extraversion/introversion has been used to describe social stereotypes through associative tasks (Andersen & Klatzky, 1987), and was also included in the character-rating study conducted by Berry and Brown (2017). An expansion of this work on the classification of literary characters is needed to optimize the categories of traits that are used in creating predictive dimensional schemes for experimental analyses of acting.

Although the current study focused exclusively on a within-subject analysis, namely the 3×3 character scheme, and did so across 12 prosodic variables, there are a number of critical between-subjects questions that should be addressed in future studies. One is the contrast between actors and nonactors, and a second is the contrast within the population of actors between different methods of getting into character. Although we are not aware of any studies that have compared actors and nonactors using a character-portrayal task, a number of studies have looked at the comparison with respect to the expression of emotion, focusing on perceivers’ ratings of authenticity between spontaneous and acted-out emotional expressions. Jürgens, Grass, Drolet, and Fischer (2015) compared examples of spontaneous emotional speech from radio interviews with reenactments of these samples for four basic emotions, as produced by both professional actors and nonactors. Spontaneous speech was rated as more authentic than the reenactments by the two groups, which didn’t differ from one another, hence failing to show an advantage for professional actors. Anikin and Lima (2018) had listeners rate the authenticity of both spontaneous and acted-out nonverbal vocalizations from published corpora of recordings. Listeners were able to distinguish the authentic versions for certain emotions more so than others, where the perceptual difference in authenticity was largest for those emotions for which acoustic differences were the largest, such as for anger, fear, and pleasure. Juslin, Laukka, and Bänziger (2017) compared spontaneous and acted-out emotions by examining a large number of samples of verbal speech (single sentences) from existing databases. Acted-out clips were rated as having higher emotional intensity than spontaneous clips, but spontaneous clips were rated as being more genuine, even after controlling for emotional intensity. Finally, Krahmer and Swerts (2008) compared acted-out emotions with those produced through an emotion-induction procedure for positive, negative, and neutral sentences and did so with both professional actors and nonactors. Facial expression, rather than vocal expression, was examined. It was found that actors produced more-extreme versions of the emotions than did the nonactors, and that the actors’ renditions were further removed from the authentic emotions produced by emotion induction than were the nonactors'. Overall, these studies suggest that, despite the intuition that actors perform realistic portrayals of emotions, actors often produce more extreme renditions of emotions that are appraised in some cases as being less authentic than spontaneous emotional expressions or the expressions of nonactors.

We suspect that a comparative prosodic analysis of character portrayal would show that actors create more-expansive renditions of characters than do nonactors, with greater use of pitch range and dynamics. A study of poetic recitation by Menninghaus, Wagner, Knoop, and Scharinger (2018) demonstrated that a (single) professional actor produced renditions of poems that were rated as being significantly more melodious than the renditions produced by several nonactors. Hence, we predict that a group analysis of actors versus nonactors would demonstrate a similar result with respect to not just pitch, but also loudness, timbre, and timing. The next between-subjects issue relates to actor training, and we can imagine two different ways of addressing this. A between-subjects study could compare a group of actors trained exclusively with gestural methods to a group of actors trained exclusively with psychological methods. However, given that many actors have a mixed training (as in the present study), a more fruitful approach might be to employ a within-subject design in which actors of mixed training are instructed to employ either gestural methods or psychological methods to get into the same set of characters. For both such studies, it would be important to use scripts that have more psychological detail than the script used in the current study, which was designed to be as neutral as possible so as to highlight character-based differences.

Finally, the modulation of vocal prosody is but one facet of the art of impersonation by actors. Facial expression and body gestures (Dael, Mortillaro, & Scherer, 2012; Ekman & Friesen, 1978; Wallbott, 1998) are equally important for conveying the impression of being a character. Ongoing motion-capture analyses from the present experiment are examining these features of character portrayal, with the ultimate aim of exploring if there are any relationships among prosody, facial expression, and body gesture, either pairwise or as a three-way relationship. For example, Thompson, Russo, and Livingstone (2010) demonstrated that when singers increased the size of sung pitch-intervals, there were corresponding facial changes, such as increases in brow height and mouth opening. Scherer and Ellgring (2007), in a study using professional actors, searched for the presence of "multimodal clusters" in the expression of basic emotions across vocal prosody, facial expression, and body gesture. They found such multimodal effects for a small number of emotions, mainly agitation, resignation, and joyful surprise, all of them emotions associated with high urgency. Acting is ultimately about creating displays of emotional expression during public performance. Hence, the relationships that we and others are observing might be components of "display rules" (Ekman & Friesen, 1969; Friberg et al., 2006) that mediate the expression of emotion in the performing arts and most likely the emotional communication that occurs during everyday social interactions as well.

Conclusions

The present study represents a first step toward establishing a psychology of acting using trained actors as the participants, as envisioned by Goldstein and Bloom (2011). The results showed that actors' prosodic productions conformed with the predictions of a two-dimensional personality scheme for characters, with much stronger effects seen for assertiveness than for cooperativeness. This supports Plato's view of acting in which actors aim to mimic the presumed personality and physical features of the portrayed characters in order to create a believable depiction of the character. The study of acting provides a means of investigating a number of issues of importance in cognitive psychology, including the multimodal expression of emotion, the impact of personality on action tendencies and social interaction, the sense of identity and the self, the nature of pretense, the manner in which consciousness can be split between the self and a portrayed character, and the means by which actors are able to embody a character through empathic mentalizing and/or gestural mimicry.

Related to this latter point, the results of the current study can be generalized to a broader perspective of role playing which views all social behavior in daily life as involving a process of playing various context-specific roles (Goffman, 1959; Landy, 1993; Shulman, 2017). In general, people modulate their vocal prosody in everyday situations to conform with their interaction partners and/or the context of the social interaction, a phenomenon that is studied in the field of interactional prosody (Auer, 2015; Couper-Kuhlen, 2014; Szczepek Reed, 2012; 2013). One indicator of this was found in the performed-self condition in the current experiment, in which the context of stage performance altered an actor's prosody compared to a conversational context. Finally, there are intermediate situations where people briefly impersonate other individuals, such as when quoting one's mother when talking about her with a friend and doing so by raising one's vocal pitch (Stec, Huiskes, & Redeker, 2015). This general process of impersonation has been referred to as "proto-acting" (Brown, 2017), and represents yet another large domain of character portrayal that can be examined in the emerging psychology of acting and role-playing.

References

Andersen, S. M., & Klatzky, R. L. (1987). Traits and social stereotypes: Levels of categorization in person perception. *Journal of Personality and Social Psychology, 53,* 235–246. http://dx.doi.org/10.1037/0022-3514.53.2.235

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *The Quarterly Journal of Experimental Psychology, 71,* 622–641.

Anolli, L., & Ciceri, R. (1997). The voice of deception: Vocal strategies of naïve and able liars. *Journal of Nonverbal Behavior, 21,* 259–284. http://dx.doi.org/10.1023/A:1024916214403

Arp, T. R., & Johnson, G. (2009). *Perrine's story & structure: An introduction to fiction* (12th ed.). Boston, MA: Cengage Learning.

Auer, P. (2015). The temporality of language in interaction: Projection and latency. In A. Deppermann & G. Günthner (Eds.), *Temporality in interaction* (pp. 27–56). Amsterdam, the Netherlands: John Benjamins Publishing Company. http://dx.doi.org/10.1075/slsi.27.01aue

Bachorowski, J. A., & Owren, M. J. (2008). Vocal expressions of emotion. *Handbook of Emotions, 3,* 196–210.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70,* 614–636. http://dx.doi.org/10.1037/0022-3514.70.3.614

Barki, H., & Hartwick, J. (2001). Interpersonal conflict and its management in information system development. *Management Information Systems Quarterly, 25,* 195–228. http://dx.doi.org/10.2307/3250929

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Belyk, M., & Brown, S. (2014). The acoustic correlates of valence depend on emotion family. *Journal of Voice, 28,* 523.e9-523.e18. http://dx.doi.org/10.1016/j.jvoice.2013.12.007

Ben-David, B. M., van Lieshout, P. H., & Leszcz, T. (2011). A resource of validated affective and neutral sentences to assess identification of emotion in spoken language after a brain injury. *Brain Injury, 25,* 206–220. http://dx.doi.org/10.3109/02699052.2010.536197

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior & Development, 36,* 847–862. http://dx.doi.org/10.1016/j.infbeh.2013.09.001

Benedetti, J. (2012). *The art of the actor: The essential history of acting from classical times to present day*. New York, NY: Routledge. http://dx.doi.org/10.4324/9780203056608

Berry, M., & Brown, S. (2017). A classification scheme for literary characters. *Psychological Thought, 10,* 288–302. http://dx.doi.org/10.5964/psyct.v10i2.237

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5,* 314–345.

Brestoff, R. (1995). *The great acting teachers and their methods*. Lyme, NH: Smith and Kraus.

Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition, 106,* 1497–1503. http://dx.doi.org/10.1016/j.cognition.2007.05.011

Brown, S. (2017). Proto-acting as a new concept: Personal mimicry and the origins of role playing. *Humanities, 6,* 43. http://dx.doi.org/10.3390/h6020043

Cangelosi, R., & Goriely, A. (2007). Component retention in principal component analysis with application to cDNA microarray data. *Biology Direct, 2,* 2. http://dx.doi.org/10.1186/1745-6150-2-2

Cartei, V., & Reby, D. (2012). Acting gay: Male actors shift the frequency components of their voices towards female values when playing homosexual characters. *Journal of Nonverbal Behavior, 36,* 79–93. http://dx.doi.org/10.1007/s10919-011-0123-4

Cohen, R. (2004). Role distance: On stage and on the merry-go-round. *Journal of Dramatic Theory and Criticism, 19,* 115–124.

Couper-Kuhlen, E. (2014). Prosody as dialogic interaction. In D. Barth-Weingarten & B. Szczepek Reed (Eds.), *Prosody and phonetics in interaction* (pp. 221–251). Mannheim, Germany: Verlag für Gesprächsforschung.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion, 12,* 1085–1101. http://dx.doi.org/10.1037/a0025737

Darwin, C. (1998). *The expression of the emotions in man and animals*. New York, NY: Oxford University Press. (Original work published 1872)

Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication, 40,* 33–60. http://dx.doi.org/10.1016/S0167-6393(02)00070-5

Doukhan, D., Rilliard, A., Rosset, S., Adda-Decker, M., & D'Alessandro, C. (2011). Prosodic analysis of a corpus of tales. In P. Cosi, R. De Mori, G. Di Fabbrizio, & R. Pieraccini (Eds.), *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 3129–3132). Belfast, Northern Ireland: ISCA Archive.

Duane, M. J. (1989). Sex differences in styles of conflict management. *Psychological Reports, 65,* 1033–1034. http://dx.doi.org/10.2466/pr0.1989.65.3.1033

Dunteman, G. H. (1989). *Principal components analysis*. Newbury Park, CA: Sage. http://dx.doi.org/10.4135/9781412985475

Dusinberre, J. (1998). Boys becoming women in Shakespeare's plays. *Shakespeare Studies, 36,* 1–28.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6,* 169–200. http://dx.doi.org/10.1080/02699939208411068

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica, 16,* 23–27. http://dx.doi.org/10.1515/semi.1976.16.1.23

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage and coding. *Semiotica, 1,* 49–98.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60,* 1497–1510. http://dx.doi.org/10.2307/1130938

Forster, E. M. (2005). *Aspects of the novel*. London, UK: Penguin Books. (Original work published 1927)

Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, 2,* 145–161. http://dx.doi.org/10.2478/v10053-008-0052-x

Gobl, C., & Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication, 40,* 189–212. http://dx.doi.org/10.1016/S0167-6393(02)00082-1

Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Anchor Books.

Goldstein, T. R., & Bloom, P. (2011). The mind on stage: Why cognitive scientists should study acting. *Trends in Cognitive Sciences, 15,* 141–142. http://dx.doi.org/10.1016/j.tics.2011.02.003

Goudbeek, M., & Scherer, K. (2010). Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America, 128,* 1322–1336. http://dx.doi.org/10.1121/1.3466853

Hirst, D. (2011). The analysis by synthesis of speech melody: From data to models. *Journal of Speech Sciences, 1,* 55–83.

Hoffman, B. (2014). *Acting and how to be good at it*. San Marino, CA: Ingenuity Press USA.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99,* 561–565. http://dx.doi.org/10.1037/0033-295X.99.3.561

Jürgens, R., Grass, A., Drolet, M., & Fischer, J. (2015). Effect of acting experience on emotion expression and recognition in voice: Non-actors provide better stimuli than expected. *Journal of Nonverbal Behavior, 39,* 195–214. http://dx.doi.org/10.1007/s10919-015-0209-5

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129,* 770–814. http://dx.doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., Laukka, P., & Bänziger, T. (2017). The mirror to our soul? Comparisons of spontaneous and posed vocal expression of emotion. *Journal of Nonverbal Behavior, 42,* 1–40. http://dx.doi.org/10.1007/s10919-017-0268-x

Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65–135). New York, NY: Oxford University Press.

Kemp, R. (2012). *Embodied acting: What neuroscience tells us about performance*. London, UK: Routledge.

Kilmann, R. H., & Thomas, K. W. (1975). Interpersonal conflict-handling behavior as reflections of Jungian personality dimensions. *Psychological Reports, 37,* 971–980. http://dx.doi.org/10.2466/pr0.1975.37.3.971

Kilmann, R. H., & Thomas, K. W. (1977). Developing a forced-choice measure of conflict-handling behavior: The "MODE" instrument. *Educational and Psychological Measurement, 37,* 309–325. http://dx.doi.org/10.1177/001316447703700204

Konijn, E. (2000). *Acting emotions*. Amsterdam, the Netherlands: Amsterdam University Press. http://dx.doi.org/10.5117/9789053564448

Krahmer, E., & Swerts, M. (2008, January). *On the role of acting skills for the collection of simulated emotional speech*. Paper presented at the Ninth Annual Conference of the International Speech Communication Association, Brisbane, Australia.

Kristeller, P. O. (1951). The modern system of the arts: A study in the history of aesthetics: Part I. *Journal of the History of Ideas, 12,* 496–527. http://dx.doi.org/10.2307/2707484

Landy, R. J. (1993). *Persona and performance: The meaning of role in drama, therapy, and everyday life*. New York, NY: Guilford Press.

Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion, 5,* 277–295. http://dx.doi.org/10.1037/1528-3542.5.3.277

Laukka, P., Juslin, P. N., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition and Emotion, 19,* 633–653. http://dx.doi.org/10.1080/02699930441000445

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development, 7,* 185–201. http://dx.doi.org/10.1080/15475441.2011.579839

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition, 129,* 362–378. http://dx.doi.org/10.1016/j.cognition.2013.07.015

Menninghaus, W., Wagner, V., Knoop, C. A., & Scharinger, M. (2018). Poetic speech melody: A crucial link between music and language. *PLoS ONE, 13,* e0205980. http://dx.doi.org/10.1371/journal.pone.0205980

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93,* 1097–1108. http://dx.doi.org/10.1121/1.405558

Owren, M. J., & Bachorowski, J. A. (2007). Measuring emotion-related vocal acoustics. In J. Coan & J. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 239–266). Oxford, UK: Oxford University Press.

Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior & Development, 14,* 415–440. http://dx.doi.org/10.1016/0163-6383(91)90031-M

Pereira, C. (2000). Dimensions of emotional meaning in speech. *Procedures ISCA Workshop on Speech and Emotion* (Vol. 1, pp. 25–28). Belfast, Northern Ireland: ISCA Archive.

Plato. (1968). *The republic* (Allan Bloom, Trans.). New York, NY: Basic Books. (Original work published 380 BCE)

Plutchik, R. (2002). *Emotions and life: Perspectives for psychology, biology, and evolution*. Washington, DC: American Psychological Association.

Rahim, M. A. (1985). A strategy of managing conflict in complex organizations. *Human Relations, 38,* 81–89. http://dx.doi.org/10.1177/001872678503800105

R Core Team. (2013). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria: Author. Retrieved from http://www.R-project.org/

Reed, B. S. (2012). Beyond the particular: Prosody and the coordination of actions. *Language and Speech, 55,* 13–34. http://dx.doi.org/10.1177/0023830911428871

Revelle, W. (2017). *Psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University. Retrieved from https://cran.r-project.org/web/packages/psych/index.html

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. http://dx.doi.org/10.1037/0033-295X.110.1.145

Schechner, R. (2013). *Performance studies: An introduction* (3rd ed.). London, UK: Routledge.

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology, 8,* 467–487. http://dx.doi.org/10.1002/ejsp.2420080405

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*(1–2), 227–256. http://dx.doi.org/10.1016/S0167-6393(02)00084-5

Scherer, K. R., & Ellgring, H. (2007). Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion, 7,* 158–171. http://dx.doi.org/10.1037/1528-3542.7.1.158

Schmidt, V. L. (2001). *45 master characters*. Georgetown, Ontario, Canada: Writer's Digest Books.

Schröder, M. (2004). Dimensional emotion representation as a basis for speech synthesis with non-extreme emotions. In E. André, L. Dybkjaer, W. Minker, & P. Heisterkamp (Eds.), *Tutorial and research workshop on affective dialogue systems* (Vol. 3068, pp. 209–220). Berlin, Germany: Springer.

Schröder, M., Cowie, R., Douglas-Cowie, E., Westerdijk, M., & Gielen, S. (2001). Acoustic correlates of emotion dimensions in view of speech synthesis. *Proceedings of Eurospeech, 1,* 87–90.

Shulman, D. (2017). *The presentation of self in contemporary social life*. Thousand Oaks, CA: SAGE Publications. http://dx.doi.org/10.4135/9781506340913

Singmann, H., Bolker, B., Westfall, H., & Aust, F. (2016). *afex: Analysis of factorial experiments* (R package version 0.16–1). https://CRAN.R-project.org/package=afex

Smith, W. R. (1971). Actor-character personality identification in a theatre production. *Empirical Research in Theatre, 1,* 29–38.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48,* 103–130. http://dx.doi.org/10.1016/S0749-596X(02)00519-3

Stanislavski, K. (1936). *An actor prepares* (E. R. Hapgood, Trans.). London, UK: Routledge.

Stec, K., Huiskes, M., & Redeker, G. (2015). Multimodal analysis of quotation in oral narratives. *Open Linguistics, 1,* 531–554. http://dx.doi.org/10.1515/opli-2015-0018

Szczepek Reed, B. B. (2013). Suprasegmentals: Prosody in conversation. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–6). Oxford, UK: Wiley-Blackwell.

Teshigawara, M. (2004). Vocally expressed emotions and stereotypes in Japanese animation: Voice qualities of the bad guys compared to those of the good guys. *Journal of the Phonetic Society, 8,* 60–76.

Thomas, K. W. (1992). Conflict and conflict management: Reflections and update. *Journal of Organizational Behavior, 13,* 265–274. http://dx.doi.org/10.1002/job.4030130307

Thompson, W. F., Russo, F. A., & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. *Psychonomic Bulletin & Review, 17,* 317–322. http://dx.doi.org/10.3758/PBR.17.3.317

Titze, I. R., & Winholtz, W. S. (1993). Effect of microphone type and placement on voice perturbation measurements. *Journal of Speech, Language, and Hearing Research, 36,* 1177–1190. http://dx.doi.org/10.1044/jshr.3606.1177

Villar, G., Arciulia, J., & Paterson, H. (2013). Vocal pitch production during lying: Beliefs about deception matter. *Psychiatry, Psychology and Law, 20,* 123–132. http://dx.doi.org/10.1080/13218719.2011.633320

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28,* 879–896. http://dx.doi.org/10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

Zarrilli, P. B. (2009). *Psychophysical acting: An intercultural approach after Stanislavski*. London, UK: Routledge.

(*Appendix follows*)

Appendix

Semantically Neutral Monologue Script

I walked into the room. A bag is in the room. It sits on the cabinet beside a clock. Digital clocks are common. There are four drawers in the cabinet. I see a rug on the floor. It looks to be expensive.

Received February 26, 2018
Revision received March 25, 2019
Accepted April 4, 2019 ■