AMERICAN PSYCHOLOGICAL ASSOCIATION
DIVISION 10
Psychology of Aesthetics, Creativity, and the Arts

ISSN: 1931-3896
2022, Vol. 16, No. 2, 272–289
http://dx.doi.org/10.1037/aca0000335

Storytelling as a Fundamental Form of Acting

Kiran Matharu, Matthew Berry, and Steven Brown
McMaster University

Acting is a process of pretending to be someone whom the actor is not. Whereas acting is often considered to be a specialized skill of trained professionals, a simple and perhaps universal form of acting occurs during oral storytelling, in which the storyteller acts out the characters of the story during the moments of dialogue and self-reflection. To examine this skill experimentally, we had both trained actors and novices read 4 fairy tales aloud. The stories contained a series of contrastive characters that spanned age, gender, and species. The major dependent variables were the vocal parameters of pitch, loudness, timbre, and speech rate. The results demonstrated that participants created distinguishable acoustic profiles for each character within a story, regardless of the story's familiarity. Monotonic trend analyses revealed the sequential changes in vocal parameters that were produced as a function of the age, gender, and species of the represented characters. Linear mixed-effects models showed a significant effect of acting training on character portrayal, with actors showing more-expansive pitch depictions than novices. We argue that portraying characters during story reading is one of the most fundamental forms of acting in human life.

Keywords: storytelling, acting, narrative, character, prosody

In *The Republic*, Plato (380 BCE/1968) argued that there are three basic forms of narrative, namely diegesis, mimesis, and their combination. In diegesis, a story is told using the voice of a narrator. In mimesis, a story is told using the voices of the characters, as seen in the portrayal of characters by theatrical and cinematic actors. Finally, in combined forms of narrative, diegesis and mimesis alternate with each other, as was seen in the recitation of Homeric epics in Plato's day. During such combined recitations, sections of narration using the narrator's voice are interleaved with segments of dialogue in which characters speak to one another, a format that is nearly universal in novels and children's stories. An effective storyteller of such combined forms of narrative has the task of portraying all of the characters in the story in a distinct manner as well as in differentiating these characters from the voice of the narrator, such as when a parent reads a bedtime story to a child. This process of transient character portrayal, referred to as proto-acting (Brown, 2017), has been proposed as being a precursor to the full-fledged acting of the dramatic actor. Compared with a dramatic actor—who typically portrays a single character over a several-hour performance—a storyteller has to portray multiple contrastive characters, often in quick alternation with one another. These characters can span not only gender and age group but also species. This is demonstrated in well-known children's stories about three bears, three little pigs, or a big bad wolf. A good storyteller therefore has to be a good actor as well. In the present study, we explore the idea that the portrayal of characters during oral storytelling is perhaps the most fundamental and universal form of acting in human life. In particular, we examine whether a storyteller differentiates the multiple characters of a story with their voice based on the characters' age, gender, and species.

Storytelling can be compared with two other forms of proto-acting, namely impressionism and ventriloquism. The impressionist portrays multiple characters but typically does so with little or no dialogue between characters. By contrast, a ventriloquist engages in constant dialogue between herself and her dummy character—requiring a constant modulation of her manner of speaking—but only for a single character. Oral storytelling combines the need to portray multiple contrastive characters found in impressionism with the need for dialogue between characters found in ventriloquism. Hence, this complex form of proto-acting demands a skilled performance from the storyteller, whether that be a professional storyteller or a parent reading a bedtime story to his child. In the present study, we explored the recitation of fairy tales as a means of addressing these phenomena.

Previous experimental work on the vocal portrayal of individuals, real or fictional, has mainly provided qualitative information about the performance but has generally lacked quantitative prosodic details (Bavelas, Gerwing, & Healing, 2014; Bretherton, 1989; Cohen, 2011; Doukhan, Rilliard, Rosset, Adda-Decker, &

This article was published Online First August 20, 2020.

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

This work was funded by a grant to Steven Brown from the Social Sciences and Humanities Research Council of Canada. We thank Carmen Tu for critical reading of the manuscript and for suggestions for improvement.

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

D'Alessandro, 2011; Sawyer, 1996; Stec, Huiskes, & Redeker, 2015, 2016; Stec, Huiskes, Wieling, & Redeker, 2017). When quantitative prosodic details are provided, they are rarely clearly linked to character types. Experimental studies of character portrayal in the literature can be grouped into three categories: the portrayal of fictional characters during oral reading, the portrayal of fictional characters during conversation, and the portrayal of real individuals during quotation in conversation. A very sparse set of studies has investigated the vocal and/or gestural portrayal of fictional characters during story reading. Doukhan et al. (2011) had a single male actor read 12 children's tales aloud and measured the vocal properties of the storytelling to develop a text-to-speech (TTS) system. The actor portrayed a total of 27 characters across the 12 stories, who varied in age, gender, and kind/species (Doukhan et al., 2011), a categorization method that the authors used in later studies (Doukhan et al., 2012, Doukhan, Rosset, Rilliard, Alessandro, & Adda-Decker 2015). Although these characters were categorized, no predictive scheme was presented for the vocal portrayal of the characters. After a preliminary vocal analysis, the characters were grouped into nine categories of prototypical characters (four female and five male) using a post hoc hierarchical cluster analysis. The vocal dependent variables included pitch, pitch variability, pitch range, loudness, loudness variability, and devoiced vowels. These variables were measured during dialogue when the characters spoke in a story. Although Doukhan et al. (2011) recorded many variables, they did not report the results for all of them for each character. Rather, only the most relevant variables were reported for the nine prototypical characters. The results showed that the youngest characters had the highest pitch and greatest loudness (but girls had the lowest loudness) and that all adult characters, regardless of gender, had comparable mean pitch and intensity. A limitation of using this study to exemplify the vocal characteristics of character portrayal is that its primary aim was to develop a TTS system, rather than to analyze character portrayal per se. However, it provides insight into the vocal portrayal of fictional characters during oral story reading.

Montaño, Alías, & Ferrer (2013), similar to Doukhan et al. (2011), examined character portrayal during story reading as a means of developing a TTS system. Similar to the earlier study, the authors looked at only a single professional storyteller. However, unlike it, they looked at only a single character, namely Harry Potter from the novel *Harry Potter and the Philosopher's Stone*, and they focused only on the vocal measurement of pitch. Because the stimuli included the spoken lines of only one character (rather than a dialogue between two characters), the lines were categorized by emotion, rather than by character. These included the basic emotions of hot anger, cold anger, joy, sadness, surprise, and fear. Compared with the control condition of neutral emotion, almost all of the basic emotions were spoken with a slower speech rate, higher pitch, and greater loudness. Sadness was the only emotion that showed a lower pitch and reduced loudness. Fear displayed a speech rate similar to neutral speech. The storyteller, without being prompted, imitated the prepubescent voice of Harry Potter.

In the more naturalistic context of pretend play, children are well known to depict characters, both real and fictional (Harris, 2000; Walton, 1999), although empirical studies of pretend play also lack quantitative details. In an analysis of the existing literature at the time, Bretherton (1989) proposed that children assume the roles of both playwright and actor when conveying stories during pretense. As playwrights, children make decisions about the plot, role distribution, and props to create the desired scene, sometimes during the course of the scene itself. As actors, children understand the concept of portraying an individual outside themselves, whether fictional or real. For example, a child may portray a father character using either their own body or a human-like toy as a surrogate. To portray these characters, children use their face, gestures, and voice. For example, they may use a wailing voice when portraying a character who is crying. Sawyer (1996) analyzed the pretend-play behaviors of 24 preschool children in a naturalistic observational study. Children portrayed characters by creating appropriate dialogue for the scene to depict the characters' emotions. They not only voiced the characters whom they embodied, but they also created voices for inanimate objects, such as toys. Children modified the prosody, lexicon, and pragmatics of their speech when depicting a character, but Sawyer (1996) did not quantitatively measure any vocal parameters in his observational study. Finally, Cohen (2011) studied preschool children in an observational study of pretend play. As with Sawyer (1996), she found examples of children modifying their voice when portraying characters. For example, when portraying adult female characters, the children used a higher-pitched voice, especially when the characters were excited. Overall, the studies mentioned thus far for both adults and children demonstrate general qualitative features of character portrayal but do not provide quantitative details about contrastive portrayals across the dramatis personae of a single story or bout of pretend play.

Before discussing the quotation of real people during conversation as a form of nonfictional character portrayal, we will mention the interesting example of fictional character portrayal that occurs during everyday conversation. When somebody wants to recount the story of a film to a friend, they may quote one or more of the film's characters during a particular scene by portraying these characters. Bavelas et al. (2014) carried out a study in which participants recounted a short scene from the 2004 movie *Shrek 2* after having just viewed the scene. Both body gesturing and vocal changes were examined. The tendency to portray a character when describing the scene occurred more often when the participants told the story to another person, compared with when they were talking on their own, arguing that character portrayal serves as a form of demonstration (Clark, 2016). For example, one participant portrayed a scene from the film in which the large ogre picked up the cat with two fingers and held the cat in front of his face. The participant used facial expressions to imitate the menacing expression of the ogre and the fearful, apologetic look of the cat, used a hand gesture with two fingers and a thumb to mime the ogre picking up the cat, and used their voice to produce a direct quotation of the fearful cat from the film. The authors' aim was to account for conversational demonstrations in general, and they did not quantify the specific acoustic vocal changes made during character portrayal. The study showed that fictional character portrayal by everyday individuals does occur in a naturalistic setting.

In a study that looked at the quotation of both fictional characters and real individuals, Blackwell, Perlman, and Fox Tree (2015) transcribed the depictions used when participants described 18 short YouTube clips to a confederate. The clips contained humans,

nonhuman animals, and personified inanimate objects, both fictional characters and real individuals. The authors found that direct quotations showed a greater number of vocal and bodily demonstrations than did indirect quotations. During these depictions, participants changed the pitch of their voice to portray the people, animals, and/or things shown in the videos. For example, a little girl was portrayed using a high pitch, whereas a dog was portrayed using a howl. Because the purpose of this study was to explore the presence of both vocal and bodily portrayals during direct or indirect quotations, it did not report a detailed vocal analysis of the produced portrayals. Wade and Clark (1993) studied direct quotations after having participants watch various scenes of videotaped dialogue one time and then recount these scenes to a listener. In principle, reported speech can be reproduced in two ways: direct quotation (He said, "Wait for me.") or indirect quotation (He told me to wait). The authors hypothesized that verbatim direct quotations would be unlikely and that people would instead paraphrase. The mean percentage of verbatim words used for indirect and direct quotations was 35% and 38%, respectively, and so most of the retellings were not verbatim. Listeners did not equate speakers' accuracy of a narrative with the accuracy of the quotations. Wade and Clark (1993) proposed that the lack of verbatim retelling in reported quotation is due to the fact that speakers are more likely to depict *how* a person speaks than what they say. For example, a participant using direct quotations may depict the sarcasm of a character using the pitch of their voice without reproducing their words verbatim. Just as with previous studies, Wade and Clark (1993) also did not provide prosodic details of any portrayal.

A final context in which people portray others is when quoting them during everyday conversations. However, unlike the examples just presented, in which people portray fictional characters, here they portray real people, such as friends or public figures like celebrities or politicians. Staging theory (Clark, 2016) proposes that speakers spontaneously depict events for interlocutors using their voice, gestures, and facial expressions. These depictions occur during storytelling when speakers stage the events of the story, whether real or fictional. There are different forms of depictions. Our focus in the present study is on embedded depictions, which use direct quotations or reported speech to reenact a real or possible conversation. It is during these direct quotations that we expect that the characters of the story should be depicted and that proto-acting should occur. Stec et al. (2015, 2016, 2017) analyzed the personal, semispontaneous narratives of speakers and annotated the various depictions that occurred during quoted speech. These depictions, including character intonation, are proposed to help portray the individuals being quoted by the speaker, whether that be themselves or another individual. Stec et al. (2015) found that intonation was used more frequently than manual gestures to portray a person. Specifically, 55.3% of all quotations used character intonation, as compared with 20.4% using manual gestures. As with the previous studies exploring quotation or character portrayal, Stec et al. (2015) mentioned the existence of vocal modulation when participants portrayed another individual but did not quantify it, merely listing it as being present or absent because it was only one of several depictions that were measured.

Overall, the literature on vocal character portrayal is limited by the fact that it does not provide quantitative detail about how speakers portray characters during storytelling. Because of these limitations, we can look to the large literature on the vocal expression of basic emotions for insights into the most relevant dependent variables for the current study on vocal prosody. Juslin and Laukka (2003), in a meta-analysis of 104 studies using five emotion categories (anger, fear, happiness, sadness, and love-tenderness), found that anger and happiness had increased speech rate, loudness, loudness variability, high-frequency energy (voice quality), mean pitch, and mean pitch variability. In contrast, sadness and tenderness showed decreased speech rate, loudness, loudness variability, high-frequency energy, mean pitch, and mean pitch variability. Using the literature on the vocal portrayal of basic emotions and the limited literature on character portrayal, we opted in the present study to use pitch, loudness, timbre, and speech rate as our dependent variables, with an emphasis on pitch.

In addition to the within-subjects variable of character, we also set out to examine the between-subjects variable of acting experience by comparing trained actors with novices when portraying fictional characters. Because actors have extensive training in the portrayal of characters, they should show more-exaggerated portrayals than nonactors. Whereas studies comparing actors and nonactors have not examined character portrayal thus far, a handful of studies have examined emotion portrayal. These studies, both laboratory and corpus analyses, have shown that the portrayals of basic emotions by nonactors tend to be appraised by raters as being more realistic and/or authentic than portrayals by trained actors (Anikin & Lima, 2018; Jürgens, Grass, Drolet, & Fischer, 2015; Krahmer & Swerts, 2008; Juslin, Laukka, & Bänziger, 2018). This might suggest that actors exaggerate their portrayals of emotions compared with nonactors. Jürgens et al. (2015) acoustically compared authentic speech recordings (obtained from a database of radio station interviews) with sentences read aloud by both actors and nonactors for the emotions of anger, fear, sadness, and happiness. Both the actors and the nonactors showed a more variable pitch profile compared with the authentic recordings. In addition, the actors showed unique differences with respect to their articulation (i.e., lower amplitude ratio, higher peak frequency, and wider bandwidth of the first formant), most likely because of their vocal training. Anikin and Lima (2018) acoustically compared seven corpora of authentic/spontaneous emotion vocalizations with acted-out versions. The vocalizations were short sounds representing the emotions of achievement, amusement, anger, disgust, fear, and pleasure. In contrast to the results of Jürgens et al. (2015), these authors found that authentic sounds included higher pitch, more-variable pitch, less-variable spectral slope, less-variable amplitude, lower harmonicity, and more-irregular temporal structure. This difference in results may be due to the fact that this study used short, nonverbal sounds, as compared with the use of full sentences by Jürgens et al. (2015). Because our study used fairy tales as stimuli, the stimuli consisted of full sentences. Finally, a small number of studies have investigated character portrayal by actors in feature films. Recorded portrayals of homosexual characters (Cartei & Reby, 2012) and Japanese anime characters (Teshigawara & Murano, 2004) have examined how actors portray characters. Male actors performed homosexual male characters with a high-pitched, female-like voice, in comparison with heterosexual male characters (Cartei & Reby, 2012). In Teshigawara and Murano (2004), the performances of anime heroes and villains by a female student were shown to be distinct from one another, as shown by unique changes in the vocal apparatus of the actress.

The principal objective of the present study was to examine the recitation of fairy tales as a fundamental form of acting, one that combines the impressionist's art of individually portraying multiple contrastive characters with the ventriloquist's art of undergoing regular alternation between characters and thus rapid and frequent vocal modification. We carried out the study with the aim of comparing two separate cohorts of participants, namely trained actors and novices, to determine how proto-acting may differ among individuals trained in character portrayal and those who are not. Participants read four fairy tales aloud while being audio recorded. We extracted the passages of narration as well as the sections of dialogue for each character from each story. We analyzed these passages for the prosodic variables of pitch, loudness, timbre, and speech rate. For each of the four stories, we created a predictive pitch scheme based on the gender, age, and species of the characters. For example, we predicted that female characters would speak with higher pitch than male characters, and that juvenile characters would speak with higher pitch than adult characters. We tested out such predictions using a monotonic trend analysis. If the trend analysis was shown to be statistically significant, it would suggest that participants did indeed modulate the pitch of their voice in the manner predicted by our a priori scheme. Next, we used linear mixed-effects models to compare the actors and the novices in between-groups analyses. We predicted that both groups would show the same overall pitch trends but that actors would demonstrate an exaggerated vocal profile compared with novices, most likely by expanding the pitch range used to portray characters. This would be in keeping with the observation that actors tend to exaggerate their prosody compared with nonactors. We also predicted that actors would show a greater consistency in their pitch profiles for individual characters over the full course of a story, whereas novices would be more variable, such that a given character's pitch might be different between the beginning and end of a story. Finally, we performed an exploratory gender analysis to examine the influence of participant gender on depictions of character gender as well as whether this relationship interacted with acting experience.

Method

Participants

Twenty undergraduate students (mean age 19.9 ± 2.7 years, age range 18–30 years) and 20 professional actors (mean age 43.3 ± 13.3 years, age range 27–63 years; mean years of acting experience 22.7 ± 12.3 years, range 4–53 years) participated in this study. The undergraduates received course credit for their participation, whereas the actors received monetary compensation. The experiment was approved by the McMaster University Research Ethics Board.

Apparatus and Stimuli

Apparatus. The experiment was conducted in a sound-attenuated room. Participants wore a head-mounted microphone connected to a Blue Icicle interface that recorded their audio into Adobe Audition (Adobe Systems Incorporated, San Jose, CA) on a desktop computer, whose monitor was turned off to reduce participant distraction. An experimenter monitored participants through a desk-mounted webcam that was connected to a computer outside the sound-attenuated room.

Stimuli. The story stimuli were presented to participants in a binder. Four classic fairy tales were used as stimuli: *Goldilocks and the Three Bears* (GTB), *Jack and the Beanstalk* (JB), *Little Red Riding Hood* (LRRH), and *Rose Red* (RR). The first three were chosen because of participant familiarity with the narratives and characters, whereas the last one was chosen to test for proto-acting with an unfamiliar story. During the story-selection process, various fairy tales were compiled from online and print sources (Foreman, 2005; Fujikawa, 2008; Gool, 1993; Guenther, 2015; The Story of Goldilocks and the Three Bears, 2016). Stories containing at least two male and two female characters were considered, whereas stories containing numerous minor characters (e.g., the two stepsisters in *Cinderella*) were excluded from consideration. For each story, the following information was recorded for each character: age, gender, species, number of dialogue sequences, and the average length of the dialogue sequence in words (see Table 1). A sequence of dialogue consisted of the text contained within a single set of quotation marks; this could contain one or more sentences.

After selecting specific fairy tales, we edited each one as follows. Each character was given approximately the same number of dialogue sequences, modulated by how regularly the character appeared in the narrative. Each story averaged approximately five dialogue sequences per character. Characters with less than two dialogue sequences were kept in the story but were omitted from the analysis. None of the dialogue sequences were interrupted by narration. For example, the following sentence was not acceptable: "This porridge is too hot," said Goldilocks, "I don't want this!" For every story, the sequences of dialogue across each character were made to average approximately 10 words per sequence. Rhyming sentences of dialogue were eliminated to avoid possible pitch effects. Each story was reduced to run approximately five minutes in length at a normal speed of recitation, or an average of 900 words per story. Table 1 contains all character information, dialogue averages across the stories, and relative-pitch predictions for each character in each story, ordered top to bottom from lowest pitch to highest pitch after the story-editing process. The characters were organized in a hierarchy of vocal pitch based on their age and gender as follows: adult (male), adult (female), child (male), senior (female), child (female), and toddler. This was further broken down by species, in which a bear, for example, would speak with a lower pitch than a wolf. In addition to sections of character speech, each fairy tale was comprised primarily of passages of narration, which were subject to analysis (see below). Lastly, participants read the standard "Rainbow Passage" (Fairbanks, 1960) to establish a baseline for their conversational pitch.

Procedure

Participants were given the fairy tale stimuli, experimental instructions, and a questionnaire several days prior to the experiment. They were able to complete the questionnaire online. The questionnaire inquired about native language, childhood reading habits, current reading habits, and experience reading to others. Upon arrival to the laboratory, participants provided their consent and completed the questionnaire in person if they had not already done so online. The experimental instructions were repeated to

Table 1
Fairy Tales and Their Characters

Title of story	Character	Species	Gender	Age group	Number of dialogue sequences	Average length of dialogue (words)	Average number of dialogues across story	Average length of dialogue across story (words)	Total number of words in story
Goldilocks and the Three Bears	Father Bear	Bear	M	Adult	5	6.2	5.8	8.6	693
	Mother Bear	Bear	F	Adult	5	7.6			
	Goldilocks	Human	F	Child	9	10.1			
	Baby Bear	Bear	M	Toddler	4	9.5			
Jack and the Beanstalk	Giant	Giant	M	Adult	7	6.0	5	10.9	996
	Strange man	Human (mystical)	M	Adult	2	12.5			
	Giant's wife	Giant	F	Adult	5	9.6			
	Mother	Human	F	Adult	5	13.0			
	Jack	Human	M	Child	6	15.3			
Little Red Riding Hood	Wolf	Wolf	M	Adult	5	15.0	4.7	11.5	979
	Woodcutter	Human	M	Adult	3	9.3			
	Mother	Human	F	Adult	5	15.4			
	Little Red Riding Hood	Human	F	Child	7	9.0			
	Grandmother	Human	F	Senior	3	11.7			
	Wolf as Grandmother	Wolf as Human	M as F	Adult	4	9.5			
Rose Red	Bear	Bear	M	Adult	5	11.4	4.8	13.3	981
	Bear as Prince	Human	M	Adult	3	22.7			
	Mother	Human	F	Adult	4	13.0			
	Rose Red	Human	F	Teen	6	8.8			
	Little girl	Witch/mystical human	F	Child	6	14.7			

Note. F = female; M = male. Characters (excluding the narrator) are listed in ascending order with regard to their predicted vocal pitch after the story-editing process.

ensure clarity. Participants were then escorted to a soundproof room to complete some vocal warm-up tasks and then the experiment. Participants were instructed to read expressively as if reading the story to a child. This instruction was important to give the participant the idea that this was a performance and that it should be tailored to an intended recipient. It needs to be pointed out that the participants were not instructed to portray the characters in the stories, only to read the stories as written. If the participant misread a sentence, they were instructed to reread the sentence from the beginning, rather than continue to the end of the sentence. The recitations were done at a self-directed pace.

During the warm-up phase, habitual speaking pitch was measured by calculating the mean vocal pitch during the recitation of the standard "Rainbow Passage" (Fairbanks, 1960). As a practice for working with the combination of dialogue and narration, participants were asked to read aloud the short story *Jerry's New Sled* (Jackson, 1969) in an expressive manner, as if they were reading to a child, but this portrayal was not measured. After completion of the warm-up phase, the experimenter left the soundproof room, and the participant began the experiment. A fairy tale was first read aloud as practice. The same fairy tale was then read again as a test trial. Only test trials were analyzed. This process was repeated for all four fairy tale stimuli, whose order of presentation was randomized.

Design

Whereas the experiment utilized a mixed-model design in which all participants read all of the stories, the stories were analyzed independently of one another in four separate analyses. There were four independent variables: character, character gender, acting experience, and participant gender. Character and character gender were within-subjects variables, whereas acting experience and participant gender were between-subjects variables. The levels of the character variable corresponded with the series of characters in each story (see Table 1). In addition to the characters, the narrator was included as a voice distinct from the participants' conversational voice. The number and length of narrator lines chosen for the analysis were matched to the average number and length of character lines within each story. For the narrator lines, the first line and last line of each story were chosen as well as three approximately equally spaced lines from the middle of the story. The selected narrator lines were required to not have character dialogue directly preceding or following them. Most stories had an average of five sequences of dialogue per character. The average length of character dialogue for all four stories was 11 words. Next, the independent variable of acting experience consisted of two groups: actors and nonactors. Lastly, gender was comprised of both the gender of the characters within the stories and the gender of the participants.

Data Analysis

Each participant's audio file was divided into individual sequences of dialogue or narration using Adobe Premiere Pro CC (Adobe Systems Incorporated). Acoustic analyses were carried out using Praat software (Boersma, 2001). An in-house script trimmed each sequence of dialogue to retain only the vocalized sections of the production. All sections of silence and noise were eliminated using a two-pass filter derived from Hirst (2011). The purpose of trimming the audio was to eliminate artificial changes to the dependent-variable measurements, for example an inflation in pitch height. The exception was the measurement of speech rate, which required the full untrimmed sample and was relatively unaffected by noise in the sample. A Praat script analyzed each untrimmed sequence of dialogue or narration to obtain a measure for duration (seconds) to calculate the speech rate. After the audio was trimmed, another Praat script analyzed various parameters of each spliced audio file, including pitch (hertz), loudness (decibels), and timbre (harmonics-to-noise ratio [HNR]). For each character and the narrator, a grand mean of pitch across all lines was calculated. The grand mean values in hertz for dialogue and narration were then converted to a cents value relative to the conversational pitch of each participant using the formula, $c = 1200 \times \log_2 (f_2/f_1)$, where f_1 is a participant's conversational pitch obtained using the Rainbow Passage. Note that 100 cents is equal to one equal-tempered semitone. For each character and the narrator, a grand mean of loudness in decibels across all lines was also calculated. The grand mean decibel values for dialogue and narration were then converted to a ratio value relative to the conversational loudness of each participant using the formula, $x = 2\hat{}(\mathrm{db}_2 - \mathrm{db}_1/10)$, where db_1 is a participant's conversational loudness obtained using the Rainbow Passage. This ratio showed how many times louder or quieter the dialogue and narration were in comparison with the conversational loudness. To measure speech rate, the number of syllables for each line was calculated using an online automated syllable counter (How Many Syllables, 2018) and was corroborated by the researchers. For each line of the story, the number of syllables was divided by the duration of the line in minutes. The speech rate values were then compared with the conversational speech rate as a percentage increase or decrease. Lastly, timbre was measured using the HNR. Line HNR values were compared with the conversational HNR using a comparative percent value, either increasing or decreasing from the conversational HNR. Before examining the data, we eliminated outliers—namely values that were two or more standard deviations from the mean—at the story level for each individual story. Problematic values often came about because of the use of creaky voice by the participant or by their decision to use whispering, despite a caveat not to do so. From a total of 5,000 sequences of character dialogue across the four dependent variables, 160 outlier data points were removed (160 of a total 20,000 data points, or $<1\%$). Instead of removing an entire sequence of dialogue, we removed only a specific outlying value within a given dependent-variable measurement to maintain the requirements for our statistical analyses.

To examine whether the produced pitch levels of the various characters both within each story and across all stories conformed to our a priori predictions, we carried out a monotonic trend analysis (Matzke et al., 2010) to look at linear trends in pitch across characters. The significance of the monotonic trend was assessed using Mann-Kendall's tau, which tests whether the dependent variable has an increasing or decreasing trend along the ordered independent variable, regardless of whether the trend is linear. Table 1 lists the characters in descending order of predicted vocal pitch. Table 2 lists all characters from all stories, ordered in pitch from left to right across columns and from top to bottom within each column, in increasing order of predicted vocal pitch. Correlations across all four vocal parameters were measured using

Table 2
Ordering of All Characters Across All Four Stories Based on Age, Gender, and Species

Level	1. Narrators	2. Lowest adult male characters	3. Secondary adult male characters	4. Adult female characters	5. Maidens and second-highest characters	6. Children and senior female characters
Character	• Narrator (GTB, JB, LRRH, RR)	• Father Bear (GTB) • Bear (RR) • Giant (JB) • Wolf (LRRH)	• Woodcutter (LRRH) • Strange man (JB) • Bear as Prince (RR)	• Giant's wife (JB) • Mother (RR) • Mother (LRRH) • Mother (JB) • Mother Bear (GTB)	• Goldilocks (GTB) • Rose Red (RR) • Little Red Riding Hood (LRRH)	• Jack (JB) • Grandmother (LRRH) • Wolf as Grandmother (LRRH) • Little girl (RR) • Baby Bear (GTB)

Note. GTB = Goldilocks and the Three Bears; JB = Jack and the Beanstalk; LRRH = Little Red Riding Hood; RR = Rose Red.

Spearman's rho. A linear mixed-effects (LME) regression model from the *lme4* package in R (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2013) was used to determine whether participants created distinguishable vocal parameters for the various characters as well as whether actors and nonactors portrayed the characters differently. A three-way LME was run for participant gender, the gender of the story characters, and acting experience.

Results

Monotonic Trend Analysis

To test the linear pitch predictions for the characters, as specified in Table 1, we ran monotonic trend analyses both for the combination of all 20 characters across the four fairy tales (which we call the FULL analysis; Figure 1) and for each of the four stories individually (see Figure 2). The results reported here are for the actors and nonactors combined; the comparison between actors and nonactors will be presented below in the context of the LME analyses. All monotonic trend analyses were highly significant using Mann-Kendall's tau statistic: FULL: $\tau = 0.246$, $p < .001$; GTB: $\tau = 0.450$, $p < .001$; JB: $\tau = 0.340$, $p < .001$; LRRH: $\tau = 0.313$, $p < .001$; and RR $\tau = 0.532$, $p < .001$. These results indicate that participants produced changes in vocal pitch that conformed with our a priori predictions based on the gender, age, and species of the characters. The narrator voice was shown to be higher than the conversational voice by more than 2 semitones, suggesting that speakers used a type of performance voice for the narrator that was distinct from their conversational voice. In addition, every character was shown to have a higher pitch than the narrator (and thus the conversational voice as well), even characters such as the Giant in JB and the Wolf in LRRH, who, as males, would be assumed to have low-pitched voices. Monotonic trend analyses were also run for the three other vocal parameters using the same ordering as for pitch, although we did not have the same level of confidence in these predictions as we did for pitch. Significant monotonic trends (all $p < .001$) were found for: loudness for FULL, GTB, LRRH, and RR; timbre for FULL and all four stories; and speech rate for GTB and JB, with a weaker trend seen for GTB than for other the stories and other variables.

Correlations

Most of the vocal parameters were found to be correlated with one another (see Table 3). This accounts, in part, for the significant monotonic-trend values found for the three exploratory variables of loudness, timbre, and speech rate when using the predictive character ordering based on pitch. All variables showed significant correlations with one another, except for the correlation between timbre and speech rate. Pitch, loudness, and timbre showed the highest correlations with one another. The strong correlation between pitch and loudness conforms with similar results found in the literature on vocal emotion (Gramming, Sundberg, Ternstrom, Leanderson, & Perkins, 1988) and with the portrayal of stock characters by professional actors (Berry & Brown, 2019).

Effect of Acting Experience

The previous analyses combined the actors and nonactors, so we now use between-subjects LME models to compare the two groups

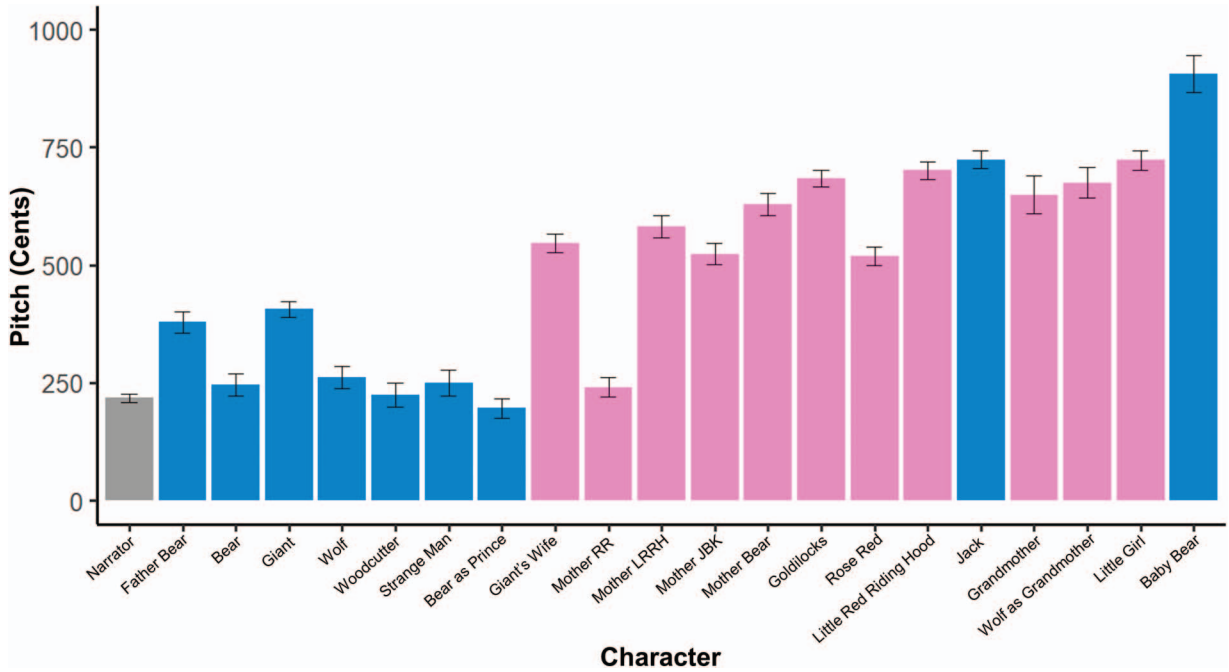

Figure 1. Mean pitch values for every character across all four stories, ordered from lowest to highest predicted pitch. Narrators are labeled in gray, male characters in blue, and female characters in pink. The narrators for all four stories have been compiled into one narrator value. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The results shown here are for the combination of the actors and nonactors. Each bar is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

of participants while including character as a within-subjects factor. Figures 3 and 4 show similar analyses for pitch to those in Figures 1 and 2, except that the results are now broken down by acting experience. Consistent with the monotonic trend analyses, there was a significant main effect of character, indicative of the differentiation of characters using pitch. This was seen for FULL, $F = 122$, $p < .001$, and for each of the four stories individually: GTB, $F = 135$, $p < .001$; JB, $F = 100$, $p < .001$; LRRH, $F = 113$, $p < .001$; and RR, $F = 139$, $p < .001$. Moreover, there was a significant main effect of acting experience. This was seen for FULL, $F = 14.8$, $p < .001$, and for each of the four stories: GTB, $F = 7.38$, $p < .01$; JB, $F = 8.09$, $p < .01$; LRRH, $F = 14.1$, $p < .001$; and RR, $F = 12.7$, $p < .001$. The actors portrayed characters with a higher pitch, on average, than the nonactors. However, this effect showed a significant interaction with character (see Figures 3 and 4). For the lower-pitched characters, the actors and nonactors portrayed the characters in a similar manner, but for the higher-pitched characters, the actors portrayed the characters with relatively higher pitch than the nonactors. In other words, the actors expanded their pitch range for the high-pitched characters more so than did the nonactors. For the highest-produced characters, this amounted to a difference of about 750 cents (7.5 semitones). This experience-by-character interaction was significant for FULL, $F = 19.4$, $p < .001$, as well as for each of the four stories: GTB, $F = 18.1$, $p < .001$; JB, $F = 17.1$, $p < .001$; LRRH, $F = 22.9$, $p < .001$; and RR, $F = 21.5$, $p < .001$.

Looking now at the exploratory variables, the vocal parameters of loudness and timbre displayed a significant main effect of character and an interaction between character and acting experience for FULL and all four stories ($p < .001$ for all analyses, except for the timbre value for the interaction between character and acting experience for RR, where $p = .016$). The patterns for both variables were similar to those for pitch, with higher-pitched characters being generally louder and having clearer timbre than lower-pitch characters. Low-pitched, dominant male characters, like the Giant and Father Bear, showed a higher pitch than expected during portrayal. However, both of these characters were also portrayed more loudly, which may have led to a higher pitch during reading.

Speech rate showed a different type of pattern (see Figure 5). The most important finding was a significant main effect of acting experience for FULL and all four of the stories ($p < .001$ for FULL, JB, and LRRH; and $p < .01$ for GTB and RR). Actors overall spoke more slowly than did nonactors, both when portraying the characters and when performing the narrator lines. In addition to the main effect of experience, there was a main effect of character on speech rate for FULL and for the four stories ($p < .001$). However, the zigzag pattern in Figure 5 indicates that it is not easy to create any generalizations regarding which characters tended to be portrayed more slowly versus more quickly.

Effects of Participant Gender and Character Gender

The last issue to be looked at statistically is an LME analysis of character gender and participant gender for pitch. Do male and female participants produce comparable portrayals of male and

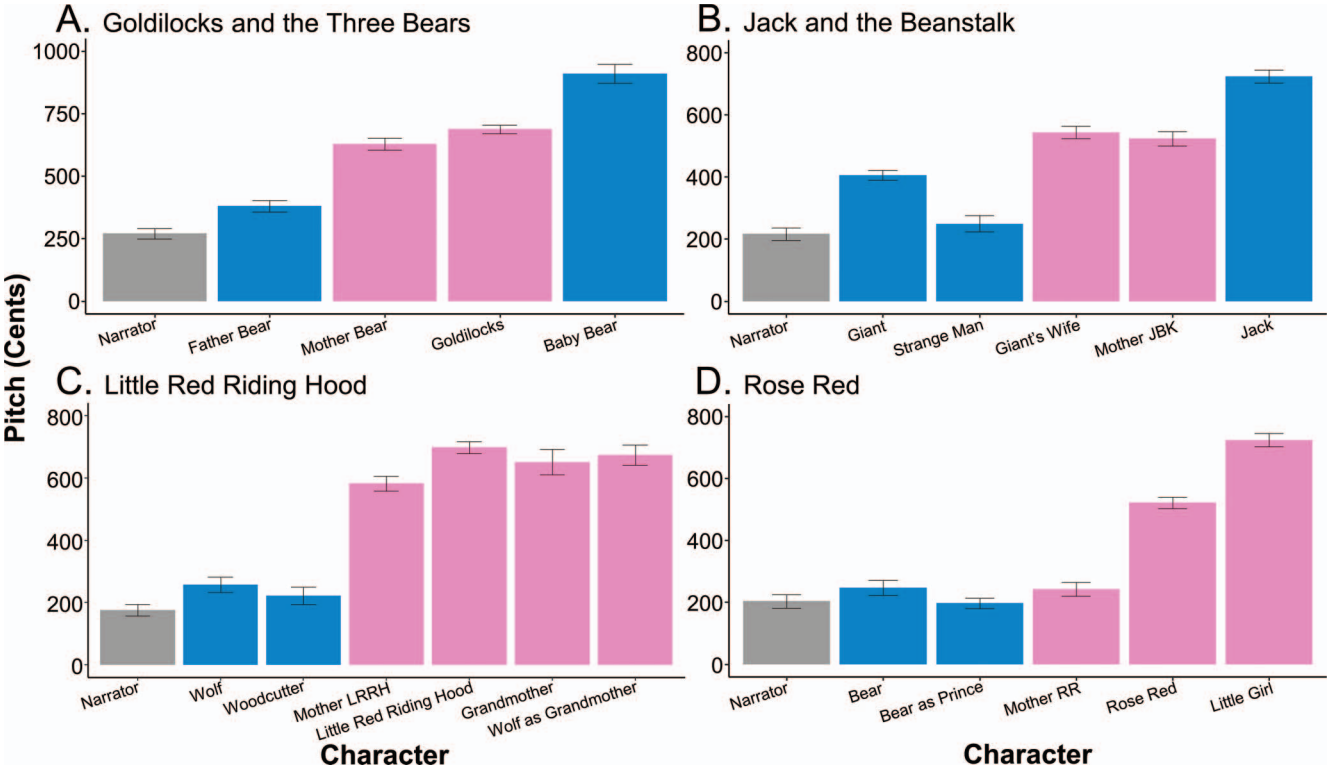

Figure 2. Mean pitch values for every character within each of the four stories, ordered from lowest to highest predicted pitch. Narrators are labeled in gray, male characters in blue, and female characters in pink. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The results shown here are for the combination of the actors and nonactors. Each bar is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

female characters? Based on our predictive pitch scheme for characters (see Table 1), we predicted that female characters would be produced higher than male characters, a prediction that was borne out for FULL, $F = 310$, $p < .001$, and three of the four individual stories: GTB, $F = 21.1$, $p < .001$; LRRH, $F = 403$, $p < .001$; and RR, $F = 194$, $p < .001$. With respect to participant gender, we had no a priori predictions. We observed no significant effect here, except for JB, for which male participants portrayed the characters higher (relative to their conversational voice) than did female participants, $F = 5.91$, $p = .02$. Other stories showed a similar, but nonsignificant, trend for males to produce higher-pitched portrayals (relative to their conversational pitch) than females. There was no interaction between character gender and participant gender. For completeness, we mention again the interaction between experience and character gender, shown in Figures 3 and 4, in which actors diverged from nonactors mainly for the high-pitched characters, which were mainly female characters, but also included a juvenile male.

Pitch Production Across the Stories

As a final analysis, we present pitch timeline graphs in Figures 6 and 7 for each story, showing the complete sequencing of the dialogue lines for each character in each story in addition to select lines of narration (see the *Method* section for a description of how the narrator lines were selected for each story). Each individual character is shown by means of a connected line running through the story. Whereas no statistics are used here, the analysis gives a qualitative indication of the extent to which participants were consistent in producing the pitch of specific characters over the course of a story. Two major findings emerge from this analysis. First, the relatively flat profile for each individual character dem-

Table 3
Correlations Between the Four Vocal Parameters

Variables	Pitch	Loudness	Timbre	Speech rate
Pitch	1			
Loudness	$\rho = .511$, $p < .001$	1		
Timbre	$\rho = .420$, $p < .001$	$\rho = .436$, $p < .001$	1	
Speech rate	$\rho = -0.167$, $p < .001$	$\rho = -0.063$, $p < .001$	$\rho = -0.003$, $p = .865$	1

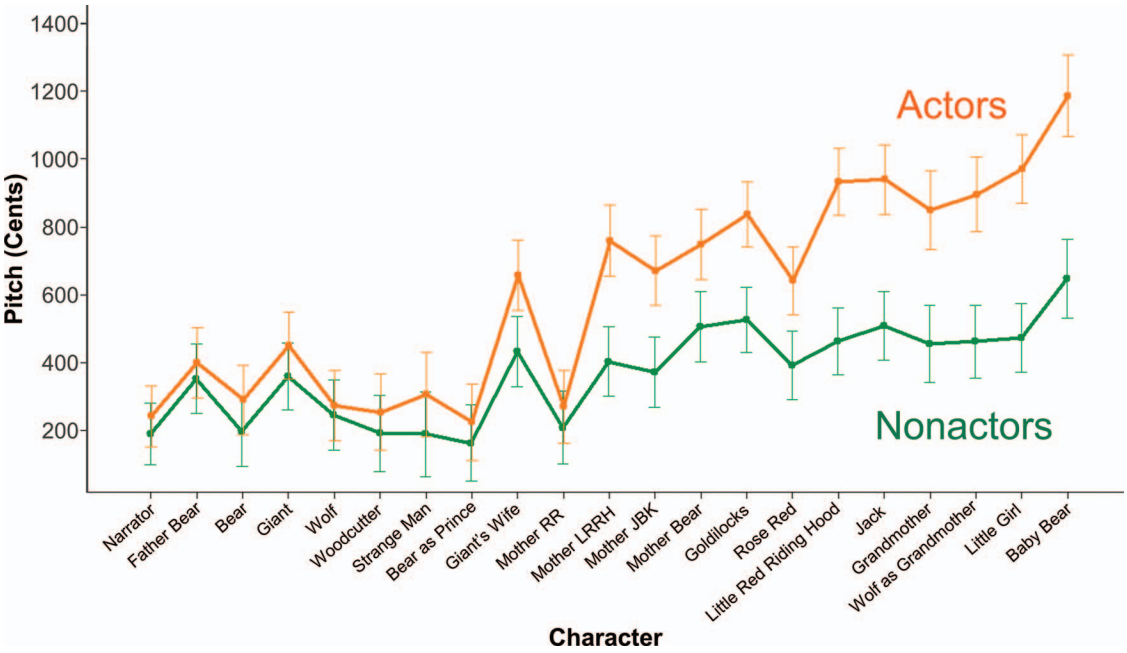

Figure 3. Mean pitch as a function of acting experience across all characters from all stories. The narrator has been compiled from all four stories. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The characters are ordered with regard to pitch predictions exactly as in Figure 1. Each value is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

onstrates that the character portrayals within each story were performed in a reliable manner, even by the nonactors. In other words, even though participants were using different pitches to represent different characters and were alternating between characters during many sections of dialogue, they tended to reliably return to the same general pitch when portraying specific characters. Exceptions were seen in high-pitched female characters like Goldilocks. Goldilocks is unique among all of the characters in the stories in that she does not physically interact with other characters or engage in dialogue with them; all of her lines are monologic. In addition, her lines are organized into groups of three in which she rejects the first two items of a category (e.g., porridge) but is pleased by the third one. Both of these factors might have contributed to the less reliable pattern for Goldilocks compared with the other characters, who tended to have much more reliable pitch patterns across the stories.

The second major finding was that actors showed a much more expansive pitch profile than the nonactors (or alternatively that the nonactors were compressed compared with the actors), allowing for a much greater differentiation of the characters from one another. This same general pattern is visible for each of the four stories, in which the characters are further spaced out from one another for the actors, compared with those same characters produced by the nonactors (Figures 6 and 7). The nonactors did not extend their vocal pitch much beyond 750 cents relative to their conversational voice, whereas actors extended their pitch almost 1500 cents above their conversational pitch. This roughly 750-cent difference represents over half an octave of pitch range used by the actors, but not the nonactors, when portraying high-pitched characters. Hence, actors more optimally differentiated characters from one another by recruiting a larger part of their vocal range for this purpose.

Discussion

In the current study, we explored whether the recitation of fairy tales can be considered as a simple form of acting. The study is the first one to quantitatively investigate the prosodic correlates of the vocal portrayal of fictional characters. We found that both nonactors and actors used systematic modulations of vocal prosody, primarily pitch, to differentiate the four to six characters in each story. The monotonic trend analyses demonstrated that participants' pitch productions conformed with our predictions about relative pitch based on the gender, age, and species of the various characters in a story. In fact, this trend remained significant in the combined set of 20 characters across all four stories as well as within each story individually. The LME analyses showed that participants created distinct depictions of all characters in all four stories. The pitch timelines revealed that participants did this quite consistently throughout a story, such that they returned to the same general pitch every time a character recurred in a story, much as a ventriloquist does when alternating between the self and the dummy character. Participants performed these portrayals despite the fact that there was no instruction to do so.

The group comparison between trained actors and nonactors demonstrated that actors produced a more exaggerated version of the nonactor prosodic results, most especially with regard to pitch and speech rate. Actors spoke more slowly than nonactors, and this effect occurred consistently across all the characters. However, the effect of acting experience on pitch interacted

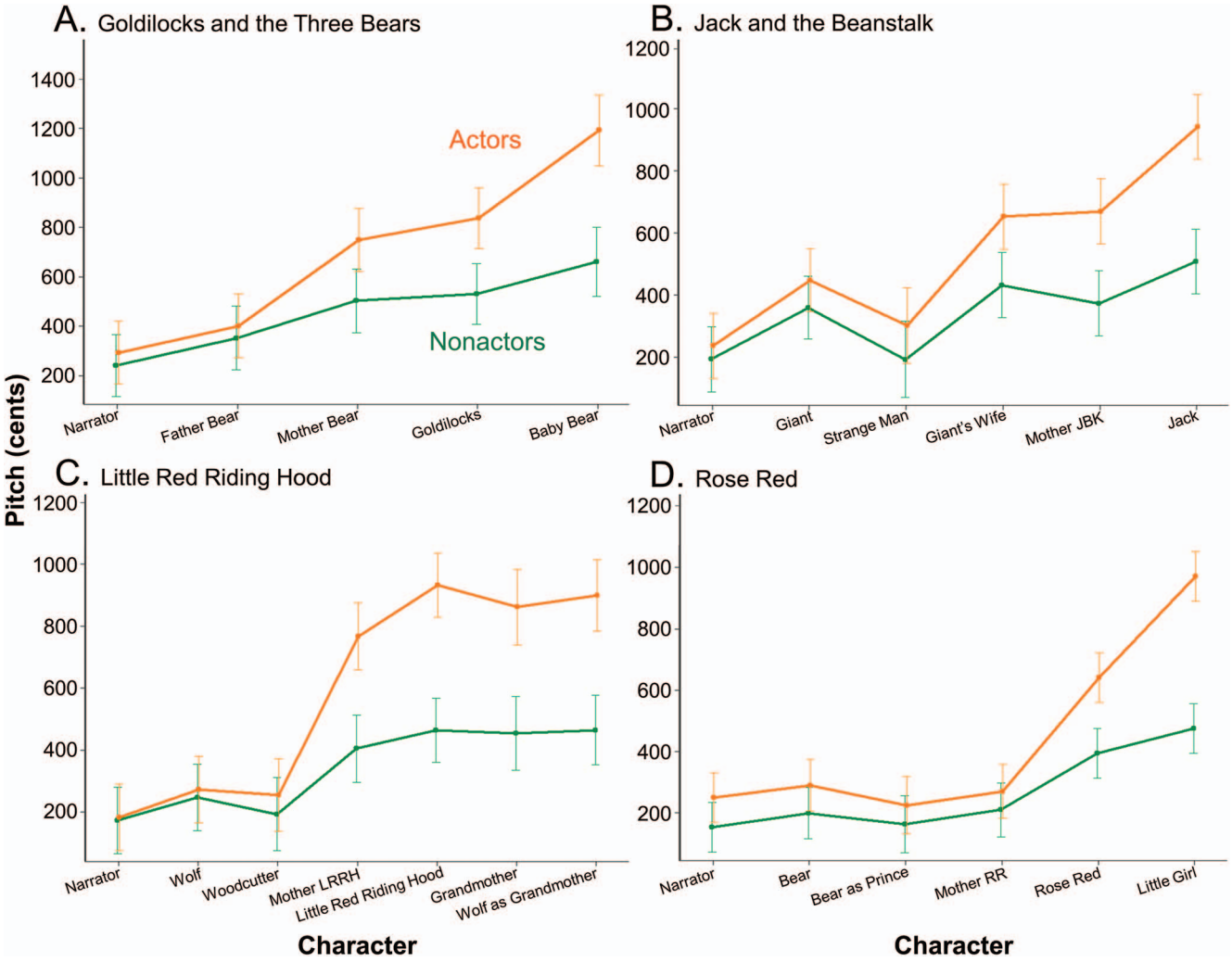

Figure 4. Mean pitch as a function of acting experience for each of the four stories. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The characters are ordered with regard to pitch predictions exactly as in Figure 2. Each value is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

with character. Whereas actors produced higher-pitched depictions than nonactors overall—using about half an octave of additional pitch range—this effect was strongly amplified for the high-pitched characters compared with the low-pitched characters. Finally, participant gender had minimal effect on the results. Because all character-portrayal measures were relative to the conversational values, existing sex differences were eliminated. Hence, this analysis shows that there was no difference in character portrayal across participant gender after removing vocal sex differences. Overall, the findings of this study support the contention that the mimesis that occurs during story reading is a fundamental form of character portrayal and thus acting. Storytelling involves systematic proto-acting.

Character Portrayal

The LME results showed that participants portrayed the characters of the stories systematically and distinctly. Both within and across all four stories, proto-acting behavior conformed with the pitch predictions established based on the age, gender, and species of the fairy tale characters. The inclusion of RR to our stimulus set ensured that character portrayal was not driven exclusively by narrative familiarity. The monotonic trend and LME analyses showed that participants portrayed the characters from RR as they did in other fairy tales. Hence, the vocal changes used to portray characters did not arise from previously heard renditions of the characters but rather from the features of the characters themselves, principally their age, gender, and species.

Previous prosodic analyses of character portrayal have reported either qualitative results alone or acoustic analyses with little quantitative detail. Doukhan et al. (2011) categorized story characters according to age, gender, and species but did not provide detailed results about how each character was vocally portrayed. Similar to Doukhan et al. (2011), we found that male bear characters were portrayed with low pitches. The characters portrayed with the lowest pitch were adult male characters. Also similar to

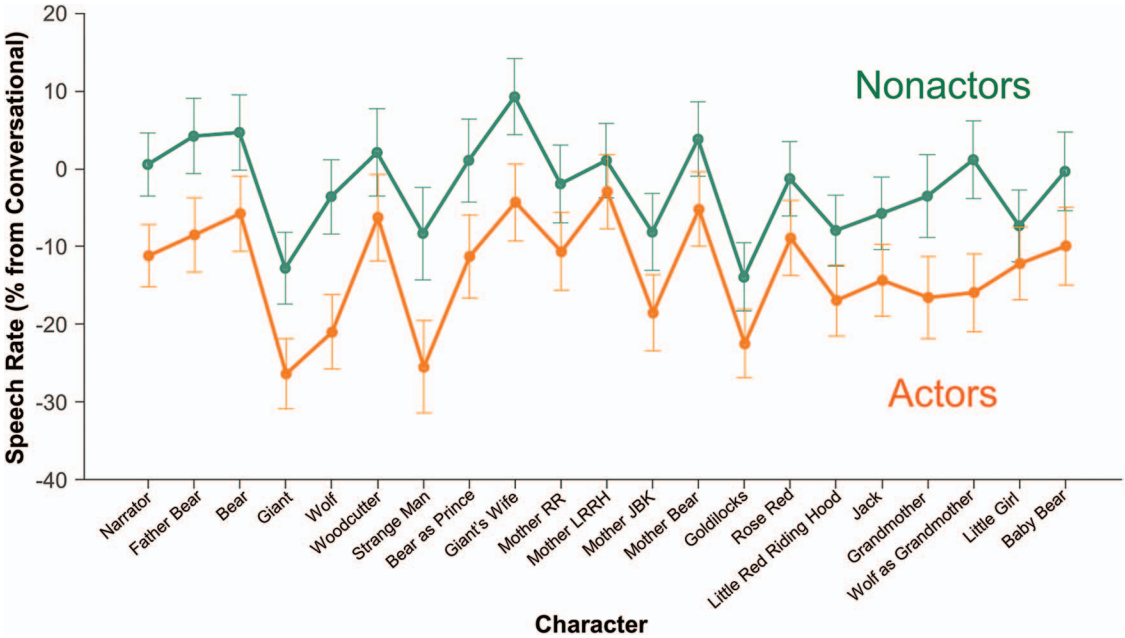

Figure 5. Effect of acting experience and character on mean speech rate across all characters from all stories. The narrator has been compiled from all four stories. The *y*-axis shows the speech rate as a percentage change from the conversational rate. Negative percentage values indicate a slower speech rate for a character compared with a participant's conversational speech rate. The characters are ordered with regard to pitch predictions exactly as in Figure 1. Each value is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

Doukhan et al. (2011), we found that child characters were generally portrayed with the highest pitches. Whereas Doukhan et al. (2011) did not specify any elderly female characters in their analysis, we found that such characters were portrayed with high pitches, similar to adult females and child characters. For example, the characters of Grandmother and Wolf as Grandmother from LRRH were portrayed in the high-pitch range.

With regard to character portrayal, we can distinguish secondary portrayal from primary portrayal. A primary portrayal would be a participant impersonating the Wolf from LRRH, whereas a secondary portrayal would be the person voicing the Wolf impersonating the Grandmother from that story. LRRH offered the sole instance of this phenomenon in our stimulus set. A comparison of the pitch results for the Wolf, Grandmother, and Wolf as Grandmother demonstrated that the Wolf as Grandmother was portrayed with a voice most similar to the Grandmother's voice and not as a male voice or as an intermediate portrayal between the two contributing characters. This demonstrates that the participant recognized that the Wolf was portraying the Grandmother and therefore adjusted his or her pitch to conform to this pretense within the story. Hence, the results suggest that a secondary character portrayal tracks the change in age, gender, and species of the secondarily portrayed character.

Vocal Registers

The monotonic trend analyses were able to reveal the overall patterns of portrayal when all the characters were organized by age, gender, and species. In addition, the timeline plots qualitatively demonstrated the reliability of each of the portrayals across the extent of a single story. The best example of this reliability was seen in the actors' portrayals of the characters from GTB, in which each character had a timeline that showed minimal deviation, even in the case of Goldilocks, who showed the most variability in pitch. Interestingly, we observed that participants' use of pitch for character portrayal did not descend below their habitual pitch, even when impersonating adult male characters or villainous characters like the Wolf in LRRH or the Giant in JB. Hence, all narration and character portrayal occurred above the habitual pitch. This phenomenon may be explained, in part, by the limited pitch space that exists below the habitual pitch, compared with that above it. Participants seemed to use loudness to overcome the pitch limitations of the lower vocal range when conveying the dominance of the male characters (data not shown). Our quantification of both the magnitude and directionality of the prosodic changes for each of the characters improves upon previous work on quotation that has merely indicated whether a prosodic change occurred but without providing quantitative acoustic details.

In examining the overall patterns of character portrayal across both the actors and the nonactors, we observed a hierarchical layering of subranges during storytelling, as shown in the summary diagram in Figure 8. First, all participants, regardless of training, utilized a higher voice than their conversational speaking voice when reciting the narrator lines, averaging about 200 cents (2 semitones) higher than the conversational voice. We refer to this below as the expression of a performance persona because the narrator serves as the presenter during storytelling. Next, the

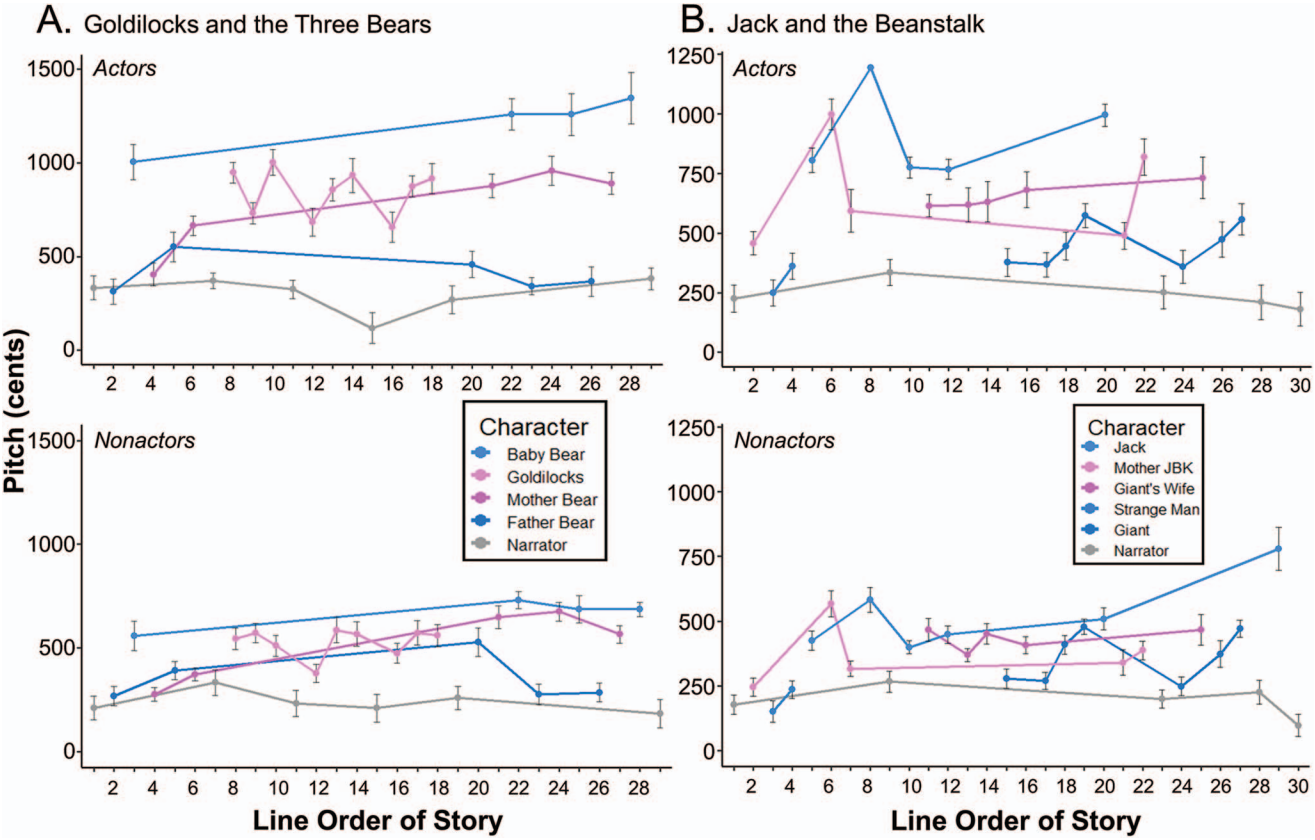

Figure 6. Mean pitch timeline graphs for the two participant groups for Goldilocks and the Three Bears and Jack and the Beanstalk. Narrators are labeled in gray, male characters in blue, and female characters in pink. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The x-axis shows the sequentially ordered lines throughout the story, in which each point applies to one line of dialogue for one character or a line by the narrator. Only lines included in the analysis are included in the timeline graph. Hence, the narrator consists of only five to six lines for each story. Each value is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

nonactors accommodated the male and female characters within a subrange that extended about 5 semitones above the narrator's pitch. This range corresponded roughly with the subrange that the actors used exclusively for the adult male characters. In other words, the nonactors' full range was contained more or less within the actors' male-specific range. Finally, the actors had an additional subrange that extended about 10 semitones above their male range for the adult female characters and the juvenile characters of both genders. This range was the main feature that distinguished the trained actors from the nonactors, the latter of whom spoke in a more compressed pitch range. Hence, as we predicted, actors produced more-expansive renditions of the characters than did the nonactors, especially for the high-pitched characters.

The observation that the narrator was performed around 2 semitones higher than the conversational voice comparably in nonactors and actors fits in with what Berry and Brown (2019) refer to as a performance persona. They described the performance persona as the presentation style that occurs "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" (p. 1421). In agreement with the results of Berry and Brown (2019), we observed here that the narrator was higher, louder, and slower than the conversational voice, indicative of a presentational style meant to be demonstrational and to capture attention. This is the prosodic style of so-called motherese that caregivers use to interact with infants (Fernald, 1985, 1989; Fernald & Simon, 1984; Grieser & Kuhl, 1988; Papoušek, Papoušek, & Symmes, 1991; Reissland, Shepherd, & Herrera, 2003; Snow, 1977; Trainor, Austin, & Desjardins, 2000; Whitehurst et al., 1988). Montaño et al. (2013) argued that storytellers use a slower speech rate to focus the attention of the audience on the performance and ensure comprehension. The performance persona is an important consideration when studying the voice. Many previous studies have used a reading voice as a control condition when measuring the pitch used to portray other people or basic emotions (Doukhan et al., 2011; Montaño et al., 2013; Revis, De Looze, & Giovanni, 2013; Stolarski,

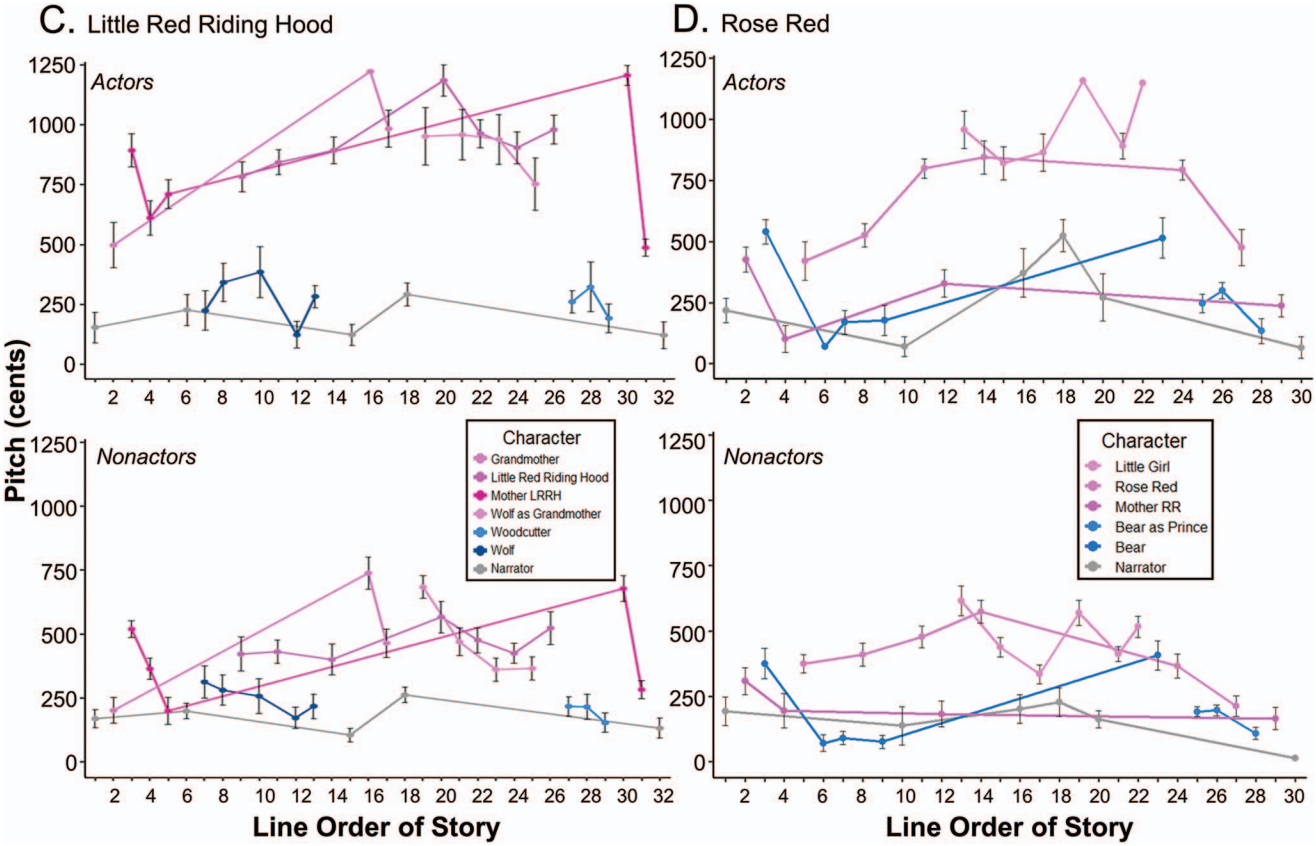

Figure 7. Mean pitch timeline graphs for the two participant groups for Little Red Riding Hood and Rose Red. Narrators are labeled in gray, male characters in blue, and female characters in pink. The y-axis shows the pitch as an increase in cents relative to the conversational pitch. The x-axis shows the sequentially ordered lines throughout the story, in which each point applies to one line of dialogue for one character or a line by the narrator. Only lines included in the analysis are included in the timeline graph. Hence, the narrator only consists of five to six lines for each story. Each value is the mean value. Error bars are standard errors of the mean. See the online article for the color version of this figure.

2015). However, a reciter's voice is not equivalent to their habitual pitch for everyday speech. Different contexts lead to different styles of reading, and this includes the vocal styles of motherese, news reading, storytelling, and the like (Doukhan et al., 2011; Snow, 1977; Theune, Meijs, Heylen, & Ordelman, 2006; Trainor et al., 2000). Hence, our demonstration that the narrator's prosody reflects the operations of a performance mode of reading indicates that future studies of oral reading need to account for the performance persona when designing their experiments.

Acting Experience

The present experiment is the first study of character portrayal to directly compare trained actors and nonactors. The results showed that even people with no formal acting training at all have a strong tendency to distinguish the narrator voice from their own conversational speech and to reliably distinguish all of the characters from one another and from the narrator in a given story. This was demonstrated in both the monotonic trend analyses and in the pitch timeline graphs. From this, we argue that character portrayal during story reading is a basic form of acting, one that does not require explicit training. A study of neutral sentence reading by nonactors showed that nonemotional speech occupies a range of about 5 semitones above a person's habitual pitch (Chow & Brown, 2018). Hence, the extension of the pitch range to 7.5 semitones (750 cents) above the habitual pitch for the nonactors in the present study is, by itself, a reflection of the impersonation of high-pitched characters. Whereas the nonactors and actors shared a pitch range when depicting the adult male characters, the major effect of actor experience on the results was to expand the pitch range of the voice when depicting the adult female and all juvenile characters. Trained actors were able to add as much as 7.5 semitones of pitch space to their range beyond the nonactors. In addition, the actors also created slower depictions than the nonactors. These findings are consistent with previous work showing that actors produced emotional vocalizations that were rated as more extreme (Krahmer & Swerts, 2008) and of higher emotional intensity (Juslin, Laukka, & Bänziger, 2018) than those of nonactors.

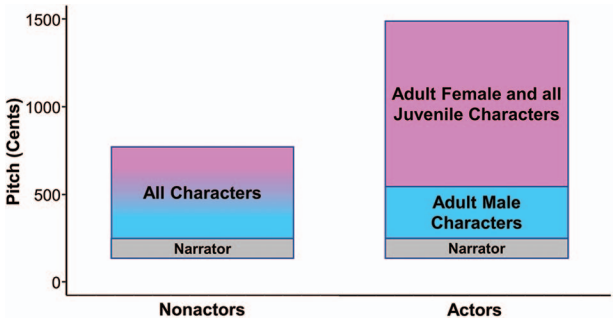

Figure 8. A schematic of the pitch results for the nonactors and actors. For both groups, the narrator was spoken roughly 2 semitones (200 cents) higher than the participant's habitual pitch, reflecting the expression of a performance persona. All of the fairy tale characters were spoken higher than the narrator's pitch for both groups. However, there was a significant divergence in pitch range between the two groups with respect to the characters. For the nonactors, all of the characters were contained in a roughly 5-semitone (500-cent) range beyond the narrator pitch, extending from around 250 cents to around 750 cents. The actors had an additional subrange extending from around 550 cents to around 1,500 cents that contained the higher-pitched characters, namely the adult females and the juvenile characters of both genders. See the online article for the color version of this figure.

Gender Differences

The gender of both the participants and the characters was relevant to this study. Because we normalized the vocal parameters of each participant to their conversational speaking voice, we were able to determine whether males and females differentially portrayed the characters of the stories, despite the fact that females speak about an octave (1200 cents) higher than males (Klatt & Klatt, 1990). Regardless of their gender, the participants used similar relative-pitch patterns in portraying fairy tale characters that varied based on age, gender, and species. Whereas the participants portrayed female characters with a higher pitch, on average, than male characters, this portrayal did not vary based on the gender of the participants. Hence, it appears that the intrinsic features of the characters were the major determinants of the pitch productions, not the features of the participants. Overall, participant gender was a null factor in our study, whereas character gender was a strong predictor of vocal prosody.

Limitations

There are several important limitations of the present study. First, because fairy tales are a short format of literature, the brevity of these stories means that there are relatively few data points for each character across a story. Using novels, as Montaño et al. (2013) did, would allow us to test character effects across a longer stretch of narrative, including examining the consistency and emotional variability of character portrayals. However, the benefit of using fairy tales was that multiple stories could be tested across a short period of time. In addition, the characters in fairy tales remain consistent across a tale, unlike the evolving characters of long-form stories. Finally, fairy tales are more commonly read aloud than are novels (typically read silently to oneself), which aids in the ecological validity of our study.

A limitation of the study was the lack of an audience during the recitation task. Bavelas et al. (2014) reported that having an active listener during a conversational task facilitated the presenter's production of depictions, especially for gestures. Other studies have shown that the rate of depictions, including vocal portrayals, is higher in a dialogue condition—whether speaking to another individual face to face or over the phone—than in a monologue condition in which the participant is recounting a story on their own (Bavelas, Chovil, Coated, & Roe, 1995, Bavelas, Coates, & Johnson, 2000, Bavelas, Gerwing, Sutton, & Prevost 2008, 2014). Although participants in our study were asked to recite the stories as if reading to a child, we might have seen an even more accentuated style of depicting characters if one or more listeners had been present to create the impression of a performance.

Because we used published stories for our study, we had minimal control over the age, gender, and species of the characters within each story. In addition, there was only one example of secondary character portrayal in the study, namely the Wolf pretending to be the Grandmother in LRRH. Furthermore, the results of some of the vocal portrayals violated our predictions based on the character's age, gender, and/or species. For example, the mother from RR was portrayed with a lower pitch than the mother characters in the other three stories. The other mothers were all similar in pitch and fell within the female subrange. However, the mother from RR was portrayed at about 250 cents above habitual, well within the male subrange. We suggest that this occurred because this mother was older than the other mothers. The character Rose Red is a maiden, whereas the other mothers have young children, hence making them younger mothers than Rose Red's mother. Another character that showed unexpected results was the Giant from JB, who was portrayed higher than expected. Because numerous studies have shown a robust positive correlation between pitch and loudness (Berry & Brown, 2019; Gramming et al., 1988), we propose that the higher pitch for the Giant may have resulted from the higher loudness used when reciting his lines of dialogue.

Another limitation of using published stories is that we had no control over the emotional features of the characters. Given the well-known effect of emotional valence and intensity on vocal parameters (see the introductory text for details), it is possible that the vocal portrayals of the characters in the present study were influenced not only by the age, gender, and species of the characters but also by emotional features that were completely dependent on the plot of the story. The only way to disentangle character and emotion for a given story is to construct stories in which the emotions of the characters are manipulated across variants of the story through changes in narration and perhaps dialogue for those characters. For example, a character of a given age, gender, and species can be described as happy in one variant and anxious in another. The impact of this difference can then be examined with respect to the vocal parameters of the storytelling.

The results of the present study can be easily extended to other contexts. A promising area of application is quotation in conversation, which is another common form of proto-acting (Brown, 2017). Instead of portraying fictional characters from well-known stories, people generally portray real individuals during improvised conversations. Future studies should be able to utilize the analysis methods of the present study to explore the depiction of people during conversational quotation. Pitch predictions can be

made based on the gender and age of the quoted individuals. For example, it is likely that a person will use a higher pitch to depict his or her mother than father, or a child compared with an adult. It is important to note that some attributions of vocal features, not least to animal characters, might be highly culture-dependent.

This work also has applications to the recent surge of interest in audiobooks. The results presented here could be used to improve the quality of oral storytelling. Further studies could look into the perception of stories read with and without the vocal portrayal of characters to see whether proto-acting during storytelling improves audience comprehension and aesthetic appreciation. If the purpose of storytelling is to convey information, then capitalizing on vocal prosody should be a natural means of increasing audience engagement in the message. For example, mothers often use a higher pitch, wider pitch range, and slower speech rate to speak to their infants (often dubbed motherese; Fernald & Simon, 1984). Infants display a preference for listening to motherese over adult-directed speech (Fernald, 1985). In addition to real-life storytellers, improvements could also be made in the development of TTS systems for storytelling speech. The systematic change in pitch that we observed when participants portrayed characters, as well as the changes in the three exploratory variables, can help to improve the verisimilitude of audiobook storytellers and TTS systems. Furthermore, acting training seems to have beneficial developmental effects for children (Goldstein & Lerner, 2018; Goldstein & Winner, 2012). After 10 months of arts training, both children and teens showed improved empathy scores after partaking in acting training, whereas training in the visual arts or music did not show the same improvement (Goldstein & Winner, 2012). In another study comparing low-socioeconomic 4-year-olds engaging in either dramatic pretend play games, block building, or story time (explicitly without enactment from the students), the children engaging in dramatic pretend-play games displayed better emotional control. They also displayed lower levels of neutral social behaviors and greater amounts of positive social behaviors (Goldstein & Lerner, 2018). In a meta-analysis of previous studies on dramatic portrayal in the classroom, children who engaged in dramatic activities displayed increased story comprehension, and this was not limited to the stories that they were enacting (Podzlony, 2000). These studies support the developmental importance of character portrayal and provide a foundation for why further explorations into proto-acting could be beneficial. Future work could explore whether stories read aloud, including character enactment, would lead to increased audience comprehension and whether children would spontaneously engage in the portrayal of characters during oral story reading, even when not instructed to.

Finally, our results have a clear application to the craft of acting and to the character portrayal that occurs during dramatic acting (Goldstein & Bloom, 2011), not to mention the related arts produced by impressionists and ventriloquists. Whereas some acting methods emphasize the psychological component of acting and thereby downplay the gestural side (Stanislavski, 1936), many classic acting methods in world cultures are firmly rooted in the gestural mimicry of character archetypes (Diderot, 1773/1883; Kemp, 2012; Schechner, 2013). Hence, we can gain insight into the nature of gestural acting through the analysis of prosodic changes during character portrayal (Berry & Brown, 2019).

Conclusions

In this first quantitative analysis of character portrayal, we found that both nonactors and actors modified their voices to portray all of the characters of a story in a distinct manner. Actors' portrayals of the characters were more exaggerated than nonactors' and utilized a larger pitch range, specifically for the high-pitched (female and juvenile) characters. Character portrayals by all participants were consistent across a given story. Storytelling involved the expression of a performance persona such that the narrator and all the characters were higher pitched, louder, clearer, and slower than a participant's normal speaking voice. There was no effect of participant gender on the results, which suggests that it is the character's features, rather than the actor's, that determine the intonation used to portray a character. These results have applications to other forms of proto-acting—such as quotation during conversation—as well as to professional acting, ventriloquism, and impressionism. Overall, these results suggest that character portrayal during storytelling is perhaps the most fundamental form of acting in human life.

References

Anikin, A., & Lima, C. F. (2018). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 71,* 622–641. http://dx.doi.org/10.1080/17470218.2016.1270976

Bates, D., Maechler, 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

Bavelas, J. B., Chovil, N., Coated, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin, 21,* 394–405. http://dx.doi.org/10.1177/0146167295214010

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology, 79,* 941–952. http://dx.doi.org/10.1037/0022-3514.79.6.941

Bavelas, J., Gerwing, J., & Healing, S. (2014). Effect of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. *Discourse Processes, 51,* 619–655. http://dx.doi.org/10.1080/0163853X.2014.883730

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language, 58,* 495–520. http://dx.doi.org/10.1016/j.jml.2007.02.004

Berry, M., & Brown, S. (2019). Acting in action: Prosodic analysis of character portrayal during acting. *Journal of Experimental Psychology: General, 148,* 1407–1425. http://dx.doi.org/10.1037/xge0000624

Blackwell, N. L., Perlman, M., & Fox Tree, J. E. (2015). Quotation as a multimodal construction. *Journal of Pragmatics, 81,* 1–7. http://dx.doi.org/10.1016/j.pragma.2015.03.004

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

Bretherton, I. (1989). Pretense: The form and function of make-believe play. *Developmental Review, 9,* 383–401. http://dx.doi.org/10.1016/0273-2297(89)90036-1

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

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

Chow, I., & Brown, S. (2018). A musical approach to speech melody. *Frontiers in Psychology, 9,* 247. http://dx.doi.org/10.3389/fpsyg.2018.00247

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review, 123,* 324–347. http://dx.doi.org/10.1037/rev0000026

Cohen, L. E. (2011). Bakhtin's carnival and pretend role play: A comparison of social contexts. *American Journal of Play, 4,* 176–203.

Diderot, D. (1883). *The paradox of acting* (H. Irving, Trans.). London, United Kingdom: Chatto & Windus (Original work published 1773).

Doukhan, D., Rilliard, A., Rosset, S., Adda-Decker, M., & D'Alessandro, C. (2011). Prosodic analysis of a corpus of tales. *Proceedings of the 12th Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 3129–3132). Florence, Italy: International Speech Communication Association (ISCA).

Doukhan, D., Rosset, S., Rilliard, A., Alessandro, C., & Adda-Decker, M. (2015). The GV-LEx corpus of tales in French. *Language Resources and Evaluation, 49,* 521–547. http://dx.doi.org/10.1007/s10579-015-9306-7

Doukhan, D., Rosset, S., Rilliard, A., Alessandro, C., & Decker, M. A. (2012). Text and speech corpora for text-to-speech synthesis of tales. *Proceedings of the 8th International Conference on Language Resources and Evaluation* (pp. 1003–1010). Istanbul, Turkey: European Language Resources Association (ELRA).

Fairbanks, G. (1960). *Voice and articulation drillbook*. Philadelphia, PA: Harper & Row.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8,* 181–195. http://dx.doi.org/10.1016/S0163-6383(85)80005-9

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

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113. http://dx.doi.org/10.1037/0012-1649.20.1.104

Foreman, M. (2005). *Classic fairy tales*. New York, NY: Sterling Publishing.

Fujikawa, G. (2008). *Fairy tales and fables*. New York, NY: Sterling Publishing.

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

Goldstein, T. R., & Lerner, M. D. (2018). Dramatic pretend play games uniquely improve emotional control in young children. *Developmental Science, 21,* e12603. http://dx.doi.org/10.1111/desc.12603

Goldstein, T. R., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development, 13,* 19–37. http://dx.doi.org/10.1080/15248372.2011.573514

Gool, V. (1993). *Van Gool's jack and the beanstalk*. Toronto: Twin Books.

Gramming, P., Sundberg, J., Ternstrom, S., Leanderson, R., & Perkins, W. H. (1988). Relationship between changes in voice pitch and loudness. *Journal of Voice, 2,* 118–126. http://dx.doi.org/10.1016/S0892-1997(88)80067-5

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology, 24,* 14–20. http://dx.doi.org/10.1037/0012-1649.24.1.14

Guenther, L. (2015). Little Red Riding Hood. Retrieved from http://www.dltk-teach.com/rhymes/littlered/story.htm.

Harris, P. L. (2000). *The work of the imagination*. Oxford, England: Blackwell.

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

How Many Syllables. (2018). *How many syllables*. Retrieved from https://www.howmanysyllables.com/

Jackson, K. (1969). *The golden book of 365 stories: A story for every day of the year*. New York, NY: Western Publishing Company.

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. (2018). 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

Kemp, R. (2012). *Embodied acting: What neuroscience tells us about performance*. London, UK: Routledge. http://dx.doi.org/10.4324/9780203126110

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87,* 820–857. http://dx.doi.org/10.1121/1.398894

Krahmer, E., & Swerts, M. (2008). On the role of acting skills for the collection of simulated emotional speech. *Proceedings of the International Conference on Spoken Language Processing (*INTERSPEECH 2008, pp. 261–264). Brisbane, Australia: International Speech Communication Association (ISCA).

Matzke, B. D., Wilson, J. E., Nuffer, L. L., Dowson, S. T., Hathaway, J. E., Hassig, N. L., . . . McKenna, S. (2010). *Visual sample plan version 6.0 user's guide*. PNNL-19915. Richland, WA: Pacific Northwest National Laboratory.

Montaño, R., Alías, F., & Ferrer, J. (2013). Prosodic analysis of storytelling discourse modes and narrative situations oriented to text-to-speech synthesis. *Proceedings of the 8th ISCA Workshop on Speech Synthesis* (pp. 171–176). Barcelona, Spain.

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

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

Podzlony, A. (2000). Strengthening verbal skills through the use of classroom drama: A clear link. *Journal of Aesthetic Education, 34,* 239–275.

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

Reissland, N., Shepherd, J., & Herrera, E. (2003). The pitch of maternal voice: A comparison of mothers suffering from depressed mood and non-depressed mothers reading books to their infants. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 44,* 255–261. http://dx.doi.org/10.1111/1469-7610.00118

Revis, J., De Looze, C., & Giovanni, A. (2013). Vocal flexibility and prosodic strategies in a professional impersonator. *Journal of Voice, 27,* 524.e23–524.e31. http://dx.doi.org/10.1016/j.jvoice.2013.01.008

Sawyer, R. K. (1996). Role voicing, gender, and age in preschool play discourse. *Discourse Processes, 22,* 289–307. http://dx.doi.org/10.1080/01638539609544976

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

Snow, C. (1977). Mothers' speech research: From input to interaction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 31–49). Cambridge, New York, NY: Cambridge University Press.

Stanislavski, C. (1936). *An actor prepares*. New York, NY: Theater Arts Books.

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

Stec, K., Huiskes, M., & Redeker, G. (2016). Multimodal quotation: Role shift practices in spoken narratives. *Journal of Pragmatics, 104,* 1–17. http://dx.doi.org/10.1016/j.pragma.2016.07.008

Stec, K., Huiskes, M., Wieling, M., & Redeker, G. (2017). Multimodal character viewpoint in quoted dialogue sequences. *Glossa, 2,* 39. http://dx.doi.org/10.5334/gjgl.255

Stolarski, L. (2015). Pitch patterns in vocal expression of "happiness" and "sadness" in the reading aloud of prose on the basis of selected audiobooks. *Research in Language, 13,* 141–162. http://dx.doi.org/10.1515/rela-2015-0016

Teshigawara, M., & Murano, E. Z. (2004). Articulatory correlates of voice qualities of good guys and bad guys in Japanese anime: An MRI study. *Proceedings of the 8th International Conference on Spoken Language Processing, INTERSPEECH 2004* (pp. 1249–1252). Jeju Island, Korea.

The Story of Goldilocks and the Three Bears. (2016). *The story of goldilocks and the three bears*. Retrieved from http://www.dltk-teach.com/rhymes/goldilocks_story.htm

Theune, M., Meijs, K., Heylen, D., & Ordelman, R. (2006). Generating expressive speech for storytelling applications. *IEEE Transactions on Audio, Speech, and Language Processing, 14,* 1137–1144. http://dx.doi.org/10.1109/TASL.2006.876129

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11,* 188–195. http://dx.doi.org/10.1111/1467-9280.00240

Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language, 32,* 805–819. http://dx.doi.org/10.1006/jmla.1993.1040

Walton, K. L. (1999). *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge, MA: Harvard University Press.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552–559. http://dx.doi.org/10.1037/0012-1649.24.4.552

Received November 17, 2019
Revision received April 11, 2020
Accepted April 28, 2020 ■