Can you hear what I feel? A validated prosodic set of angry, happy, and neutral Italian pseudowords

Emanuele Preti · Chiara Suttora · Juliette Richetin

© Psychonomic Society, Inc. 2015

Abstract This contribution aims to establish a set of validated vocal Italian pseudowords that convey three emotional tones (angry, happy, and neutral) for prosodic emotional processing research. We elaborated the materials by following a series of specific steps. First, we tested the valence of a set of written pseudowords generated by specific software. Two Italian actors (male and female) then recorded the resulting subset of linguistically legal and neutral pseudowords in three emotional tones. Finally, on the basis of the results of independent ratings of emotional intensity, we selected a set of 30 audio stimuli expressed in each of the three different emotions. Acoustic analyses indicated that the prosodic indexes of fundamental frequency, vocal intensity, and speech rate anchored individual perceptions of the emotions expressed. Finally, the acoustic profile of the set of emotional stimuli confirmed previous findings. The happy tone stimuli showed high $f0$ values, high intensity, high pitch variability, and a faster speech rate. The angry tone stimuli were also characterized by high $f0$ and intensity, but by relatively smaller pitch variability and a lower speech rate. This last profile echoes the description of "cold anger." This new set of prosodic emotion stimuli will constitute a useful resource for future research that requires emotional prosody materials. It could be used both for Italian and for cross-language studies.

Keywords Stimulus set · Pseudowords · Emotions · Prosody · Acoustic profile

Introduction

Social interactions among individuals are based primarily on the correct encoding and decoding of emotions. The literature on emotion processing has traditionally favored the study of facial emotion recognition (Ekman & Friesen, 1976; Ekman, Friesen, & Hager, 2002). However, individuals use multiple cues from different communicative channels (e.g., facial expression, body posture, and speech) to succeed in these processes (Scherer & Scherer, 2011). Moreover, difficulties in integrating multichannel emotional information can lead to impairments in empathic abilities and social cognition (Minzenberg, Poole, & Vinogradov, 2006; Preißler, Dziobek, Ritter, Heekeren, & Roepke, 2010). For this reason, researchers have recently started investigating what is now defined as *emotional prosody*.

Depending on how they are pronounced, words and utterances can carry different emotional meanings, over and above their semantic content (Banse & Scherer, 1996). Respiration, phonation, and articulation vary with the emotional state of the speaker (Scherer, 1989). Thus, the following acoustic parameters are crucial in the encoding and decoding of different emotions: the level, range, and contour of fundamental frequency ($f0$, in hertz); voice intensity (in decibels), and temporal phenomena (tempo, duration, pauses). Two reviews of numerous studies concerning the associations between the vocal expression of emotions and specific acoustic profiles (Banse & Scherer, 1996; Juslin & Laukka, 2003) showed that different combinations of speech rate, pitch level and variation, and intensity convey different emotions. These extensive reviews allow for defining prototypical profiles for each emotion. For instance, the acoustic cues that lead to identifying anger are a fast speech rate, high voice intensity levels and variability, high pitch level and variability. Happiness shares most of these characteristics, but it is defined by a medium-high voice intensity and a less intensity variability. On the basis of these acoustic features of emotions, several studies have investigated how behavioral (Hawk, van Kleef, Fischer, & van der Schalk, 2009; Paulmann, Pell, & Kotz, 2008) and neural mechanisms (Jessen & Kotz, 2011; Peelen, Atkinson, & Vuilleumier, 2010) underlie prosodic emotion processing. Furthermore, researchers have started to shed light on

E. Preti (✉) · C. Suttora · J. Richetin
Department of Psychology, University of Milan-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy
e-mail: emanuele.preti@unimib.it

🖄 Springer

prosodic emotion processing in clinical populations, considering both brain-injured patients (Dellacherie, Hasboun, Baulac, Belin, & Samson, 2011) and psychiatric conditions (Jones et al., 2011; Minzenberg et al., 2006).

To conduct such research, one needs stimuli with controlled characteristics. In fact, the development of a set of vocal stimuli has to take into account several issues. First, the acoustic features of emotional prosody are shaped by the linguistic properties of a specific language (Pell, 2001). Furthermore, lexico-semantic cues pertaining to emotions need to be controlled and isolated to understand the specific role of prosody in the expression of the emotional content. In this regard, some studies have tried to reduce the semantic information of speech through filtering procedures that maintain only the supra-segmental features of vocal stimuli (e.g., Kotz et al., 2003). Other studies have addressed this issue by developing and validating sets of pseudowords or pseudoutterances (Bach et al., 2008; Pell, 2002). Pseudowords are defined as legal nonwords that conform to the orthographic and phonological patterns of a given language. Pseudoutterances are composed of pseudowords embedded in a legal utterance. This method allows for reducing substantially the lexical and semantic properties of such stimuli while conveying language qualities, phonetic-segmental, and suprasegmental features of speech comparable to the real language.

Researchers have developed sets of stimuli with such characteristics for most Indo-European languages. Burkhardt et al., (2005) developed stimuli in German, Castro and Lima (2010) developed a Portuguese set of pseudoutterances, and Juslin and Laukka (2001) developed a series of Swedish emotional sentences. One of the most impressive databases is the multilanguage (Spanish, German, and English) stimuli set of pseudoutterances created by Pell and colleagues (Pell et al., 2009). For non-Indo-European languages, sets of emotional pseudowords/pseudoutterances are available in Arabic and Mandarin Chinese (Liu & Pell, 2012). These studies follow similar procedures, generally consisting of pseudoword/ pseudoutterance stimuli construction, recording, validation from a pool of participants, and acoustic analyses. However, in some of these contributions, there are a couple of important methodological issues concerning the set of stimuli.

The first issue concerns the generation of proper linguistically legal pseudowords. Stimuli should follow the rules of the language (language-likeness). For example, they should contain a plausible suffix (e.g., "i" or "o" for Italian pseudowords) or a pronounceable sequence of letters (e.g., "wczo" would look like a word for a Polish person but not for an Italian). As was highlighted by Keuleers and Brysbaert (2010), the predominant procedure for the creation of pseudowords consists of changing one or two letters from a list of legal words and of evaluating the language likeness only from the researchers' judgment. With the exception of Liu and Pell (2012), who asked participants to rate the language-likeness of a list of pseudoutterances, the elaboration of the stimuli has not been a fully controlled process. The second issue has to do with the guarantee of the neutrality of the pseudowords before pronunciation. Although the pseudowords do not exist, they could elicit valence. This is specially the case when the pseudowords are created by changing one or two letters of an existing word. Because of the similarity of the existing word, a pseudoword could also carry a similar valence. This issue can be addressed only by means of valence ratings of the pseudowords.

The main objective of this study was to develop a set of Italian pseudowords with angry, happy, and neutral intonations. To our knowledge, validated prosodic stimuli— pseudowords or pseudoutterances—are not yet available in Italian. The few studies addressing the emotional processing of prosodic stimuli have generally used prosodic subtests of emotion recognition batteries (such as the Comprehensive Affect Testing System; Schaffer, Wisniewski, Dahdah, & Froming, 2009), in which actors produce sentences with different emotional tones (Ariatti, Benuzzi, & Nichelli, 2008; Castagna et al., 2013). Although most of the above-mentioned sets of stimuli covered a wide range of basic emotions, we chose to focus on angry, happy, and neutral emotional prosody. Our choice was motivated by two main elements. First, in research on facial emotion information processing (e.g., Hagenhoff et al., 2013; Kirsh & Mounts, 2007), several neuroimaging and electrophysiological studies on emotional prosody have examined only one positive and one negative emotion (e.g., Kotz & Paulmann, 2007; Kotz et al., 2003; Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Schirmer, Kotz, & Friederici, 2002). Second, anger and happiness have been used frequently to operationalize negative and positive emotions, particularly when comparing one with the other or in terms of neutral emotion (i.e., Kotz et al., 2003).

In developing such a set of stimuli, we aimed to control the above-mentioned methodological issues. First, to devise a set of pseudowords formally controlled for their language-likeness, we used a specific software that controls for the subsyllabic structure and transition frequencies of the specific language (Wuggy; Keuleers, & Brysbaert, 2010). Although the procedure of pseudoword creation substantially reduces the lexical–semantic meaning of the stimulus, it does not warrant a neutral valence. We thus conducted a pretest to select neutral stimuli in their written version. Then, a man and a woman recorded the most neutral stimuli in three different emotional tones. Independent judges rated the audio stimuli for their emotional intensity to allow us to select the most prototypical audio stimuli in each emotional tone. Another sample rated the valence and the arousal of each selected audio stimuli. Finally, we examined with statistical analyses whether each emotional set of stimuli was characterized by specific acoustic profiles.

Pseudoword stimulus generation and neutral valence

Method

We used the Wuggy software (Keuleers & Brysbaert, 2010) to generate pseudowords. By means of a particular algorithm, and starting from a given list of legal syllabified words, this software generates pseudowords that obey a specific language's phonotactic constraints and transition frequencies. This procedure warrants the generation of legal pseudowords conformed to the orthographic and phonological patterns of a given language. Using the Italian submodule of Wuggy, we obtained 150 trisyllabic pseudowords. We selected 100 of these pseudowords on the basis of their ease of pronunciation for Italians. We then aimed to select a set of stimuli that would have the most neutral content in valence. For this purpose, 30 university Italian native speaker students (15 men, 15 women; M age = 22.33, SD = 2.72, range 19–27 years) evaluated the 100 pseudowords in their written form on 9-point scales that ranged from 1 (*very negative*) to 9 (*very positive*). We created four versions of the evaluation sheet with different random presentation orders of the pseudowords.

Results

We performed a series of one-sample t tests (test value = 5) on the mean valence of each of the 100 written pseudowords. We chose the 40 most neutral pseudowords (see Table 3 in the Appendix for the list of 40 pseudowords). The selected set had a mean valence of 4.95 (SD = 0.79), not different from the midpoint of the scale, $t(29) = -0.35$, $p = .731$.

Audio tokens, emotion perception, and prosodic properties

With this study, we had four main objectives. First, we aimed to examine the relationship between emotion perception and the acoustic properties of the audio versions of pseudowords recorded with three different emotional prosodies or tones (i.e., happiness, anger, and neutral). Moreover, on the basis of the independent emotion ratings, we intended to select the best audio token for each of the three emotional tones of the 30 pseudowords. We also wished to examine the perceived emotion, valence, and arousal ratings of each emotion tone set. Finally, and most importantly, we aimed to establish the acoustic profiles of the three emotional categories of the selected audio stimuli.

Method

Two professional actors (a man and a woman) recorded the 40 selected pseudowords. The male actor had a baritone-like voice (mean $f0$ of all recordings = 144 Hz), and the female actor had a mezzo soprano-like voice (mean $f0$ of all recordings = 215 Hz). Each of them recorded 20 randomly selected pseudowords in three different emotional tones (neutral, happy, angry) with five recordings for each (300 tokens for each actor). The stimuli were recorded in a sound-insulated booth at the Media Laboratory at the University of Milano-Bicocca, using Pro Tools version 10.3.7 software. A high-quality Neumann Tlm 102 microphone was connected to an Apple Macintosh Mac Pro 3.1 QuadCore Intel Xeon computer with a Focusrite Saffire Pro24 DSP audio interface and a SPL Track One voice channel. Digitization was performed at a 44.100-kHz sampling rate and a 24-bit resolution. The peak amplitudes of all pseudowords were normalized to mitigate gross differences in perceived loudness. Pseudowords conveying the three different emotional meanings were recorded in separate blocks. Two researchers monitored the recording procedure, giving cues to the actors about the target emotions. Then, three judges first selected the 30 audio pseudowords with the best pronunciation (15 for the male actor, 15 for the female actor). Then, for each pseudoword in each of the three emotional tones (neutral, happy, and angry), they retained the three best audio tokens in terms of acoustic quality and expressed emotion.

After this preselection, we gathered the intensity ratings of the three different emotions and acoustic properties information for each of the 270 audio tokens (30 pseudowords × 3 emotional tones × 3 recordings). Forty-six native Italian-speaking university students (23 men, 23 women, M age = 22.38, SD = 3.06, range 19–31 years) indicated the intensities of the 270 audio tokens on the happy, neutral, and angry dimensions. Half of the participants first rated all the 270 tokens on the happy dimension, then the neutral dimension, and then the angry dimension. The other half started with the angry dimension, then the neutral dimension, and finally the happy dimension. Within each dimension, the tokens were presented in a random order. The participants were equipped with headphones and set the volume at a comfortable level at the beginning of the procedure. After listening to each token, participants indicated how angry/happy/neutral the pronunciation of the pseudoword was on 21-point scales, from 0 (*not at all*) to 20 (*very much*), with 10 indicating *intermediate*. The interrater reliability was .99.

For the acoustic analysis, we considered different prosodic indexes by using Praat software (Boersma, 2001). On the basis of previous studies of emotional prosody (for a review, see Juslin & Laukka, 2003), we considered the mean fundamental frequency (mean $f0$, in hertz), pitch variation (mean $f0$ excursion, in semitones[1]), mean intensity (in decibels), and speech rate (duration[2]).

---

[1] The $f0$ excursion in semitones was preferred to $f0$min, $f0$max, and $\Delta f0$ in order to control for gender $f0$ natural differences (Henton, 1989).
[2] Speech rate was calculated as duration (in seconds), given the trisyllabic structure of the stimuli.

Table 1 Correlations between acoustic indexes and mean perceived emotion ratings of the 270 pseudowords

	1	2	3	4	5	6	7
1. $f0$	1						
2. Intensity	.55***	1					
3. Duration	−.58***	−.25***	1				
4. $f0$ semitones	.09	.17**	.06	1			
5. Perceived anger	−.18**	.04	.43***	−.05	1		
6. Perceived neutrality	−.42***	−.43***	−.04	−.34***	−.36***	1	
7. Perceived happiness	.38***	.18**	−.41***	.21**	−.91***	−.03	1

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

Results and discussion

Perceived emotion intensity and acoustic properties of all 270 tokens

We first collapsed the ratings of the 46 participants to produce mean ratings for each of the 270 tokens; only these mean ratings were used in our analysis. We then examined the relationships between the mean ratings and the acoustic properties of the 270 stimuli in two steps. We computed the correlations between all measures to determine whether some acoustic indexes were related to the angry, happy, and neutral ratings, respectively. Then, we ran a series of regression analyses to examine which acoustic index was most predictive of the three different ratings. For the regression analyses, and due to the dependency between some indexes, we only considered $f0$, intensity, and duration. We ran these analyses entering the gender of the actor as a covariate, because previous research had demonstrated that men and women may differ in terms of their ability to impersonate emotions (e.g., Bonebright, Thompson, & Leger, 1996).

Perceived anger As is shown in Table 1, the perception of anger correlated with all indexes with the exception of intensity and $f0$ excursion in semitones. When we entered all three indexes ($f0$, intensity, and duration) in a regression analysis for predicting the mean ratings of perceived anger, the model explained 20 % of the variance. Duration ($\beta = .47$, $p < .001$) and intensity ($\beta = .14$, $p = .031$) were significant predictors, whereas $f0$ was not ($\beta = .02$, $p = .832$). When we added gender in a second step, the model explained 28.2 % of the variance ($\Delta R^2 = .08$), $F(1, 265) = 28.96$, $p < .001$. Duration ($\beta = .57$, $p < .001$), intensity ($\beta = .27$, $p < .001$), $f0$ ($\beta = −.27$, $p = .003$), and gender ($\beta = .42$, $p < .001$) were all significant predictors.[3] Note that

when gender was entered alone in a first step, it did not predict the perceived anger ($p = .408$).

Perceived happiness Ratings of happiness correlated with all acoustic indexes (see Table 1). When all three indexes were entered in a regression analysis, the model explained 20.1 % of the variance. Duration ($\beta = −.29$, $p < .001$) and $f0$ ($\beta = .22$, $p = .005$) were significant predictors, whereas intensity ($\beta = −.01$, $p = .883$) was not. When gender was added in a second step, the model explained 40.4 % of the variance ($\Delta R^2 = .20$), $F(1, 265) = 90.24$, $p < .001$. Duration ($\beta = −.45$, $p < .001$), $f0$ ($\beta = .68$, $p < .001$), intensity ($\beta = −.22$, $p < .001$), and gender ($\beta = −.67$, $p < .001$) were all significant predictors.[4] When gender was entered alone in a first step, it was not a significant predictor of perceived happiness ($p = .934$).

Perceived neutrality Perceived neutrality correlated with all indexes with the exception of duration (see Table 1). When all three indexes were entered in a regression analysis, the model explained 34.1 % of the variance. Duration ($\beta = −.40$, $p < .001$), intensity ($\beta = −.25$, $p < .001$), and $f0$ ($\beta = −.52$, $p < .001$) were significant predictors. When gender was added in a second step, the model explained 48.2 % of the variance ($\Delta R^2 = .14$), $F(1, 265) = 72.27$, $p < .001$. Duration ($\beta = −.27$, $p < .001$), $f0$ ($\beta = −.90$, $p < .001$), and gender ($\beta = .56$, $p < .001$) were all significant predictors, and intensity ($\beta = −.07$, $p = .211$) was no longer a significant predictor.[5] When gender was

---

[3] When we entered $f0$ excursion in semitones, instead of $f0$, along with intensity and duration for predicting perceived anger, the model explained 20.6 % of the variance. Duration ($\beta = .48$, $p < .001$), intensity ($\beta = .17$, $p = .003$), and $f0$ semitones ($\beta = −.11$, $p = .044$) were significant predictors. When gender was added in a second step, the model explained 24.7 % of the variance ($\Delta R^2 = .04$), $F(1, 265) = 15.34$, $p < .001$. Duration ($\beta = .62$, $p < .001$), intensity ($\beta = .16$, $p = .005$), and gender ($\beta = .27$, $p < .001$) were significant predictors, whereas $f0$ semitones ($\beta = −.02$, $p = .706$) was not.

[4] When we entered $f0$ excursion in semitones, instead of $f0$, the model explained 22.6 % of the variance. Duration ($\beta = −.42$, $p < .001$) and $f0$ semitones ($\beta = .23$, $p < .001$) were significant predictors, whereas intensity ($\beta = .04$, $p = .480$) was not. When gender was added in a second step, the model explained 26.8 % of the variance ($\Delta R^2 = .04$), $F(1, 265) = 15.18$, $p < .001$. Duration ($\beta = −.56$, $p < .001$), $f0$ semitones ($\beta = .14$, $p = .02$), and gender ($\beta = −.27$, $p < .001$) were significant predictors, whereas intensity ($\beta = .05$, $p = .337$) was not.

[5] When we entered $f0$ excursion in semitones, instead of $f0$, the model explained 27.5 % of the variance. Duration ($\beta = −.13$, $p = .016$), intensity ($\beta = −.42$, $p < .001$), and $f0$ semitones ($\beta = −.26$, $p < .001$) were significant predictors. When gender was added in a second step, the model explained 27.5 % of the variance ($\Delta R^2 = .00$, $F(1, 265) = .03$, $p = .860$). Duration ($\beta = −.14$, $p = .037$), $f0$ semitones ($\beta = −.27$, $p < .001$), and intensity ($\beta = −.42$, $p < .001$) were significant predictors, whereas gender ($\beta = −.01$, $p = .860$) was no longer a significant predictor.

entered alone in a first step, it did not predict the perceived neutrality, although it showed tendency toward significance (p = .076).

In line with results of previous studies (Bachorowski & Owren, 1995; Castro & Lima, 2010; Liu & Pell, 2012; Pell, Monetta, et al., 2009), our analyses confirmed the central role of some acoustic parameters in determining the perception of emotions. Notably, the acoustic parameters were significant predictors also controlling for gender. Considering pitch variability (f0 semitones, notes 3–5), our results report a role of this parameter in predicting happiness but not anger. This result is not in line with the canonical acoustic profiles of anger (Juslin & Laukka, 2003). Nonetheless, as was noted by Banse and Scherer (1996), anger can be connoted as being either "hot" or "cold". In this context, high f0 variability can be considered a specific feature of hot but not of cold anger. We could thus conclude that our stimuli are much more representative of cold anger. Taken together, these results underline that speech rate, vocal intensity, and f0 represent essential acoustic cues in conveying emotional valence.

Selection of 30 pseudowords for each emotional tone

Considering the 46 participants' ratings, we then aimed to select the most representative token of each emotional tone (angriness vs. neutrality vs. happiness) for each of the 30 pseudowords. We performed a series of repeated measures analyses of variance (ANOVAs) on the ratings of the three different tokens for each emotion, with post-hoc analyses to choose the most representative. Then, we performed another series of repeated measures ANOVAs on the ratings of the tokens chosen for each emotion, with post-hoc analyses to ascertain, for example, whether the pronunciation of the angry token of "andori" was significantly perceived as being angry rather than happy or neutral. Of the 30 pseudowords, two tokens of the neutral category pronounced by the man did not show satisfactory neutrality ratings. First, the mean neutrality rating of the selected neutral token of the pseudoword "psilumbo" (M = 7.87, SD = 4.12) did not differ significantly from the mean happiness rating (M = 7.87, SD = 4.12, p = .258). Second, for the selected neutral token of "ervetto," although the neutrality rating (M = 11.24, SD = 5.87) did differ significantly from the happiness (M = 7.22, SD = 3.39) and anger (M = 7.89, SD = 4.51) ratings, it was weaker than the mean neutrality ratings of the other 13 neutral tokens (M = 14.70).

To replace these two pseudowords, we ran an additional study on the remaining ten neutral pseudowords. Thirty-two native Italian-speaking university students (17 men, 15 women, M age = 23.13, SD = 4.05, range: 20–40 years) indicated the intensity of the 90 tokens pronounced by the same man (10

pseudowords × 3 emotions × 3 recordings) according to the happy, neutral, and angry dimensions (interrater reliability, α = .98). On the basis of the mean ratings, we selected the pseudowords "rantaglia" and "zellani." Please refer to the Appendix, Tables 4 and 5, for the mean emotional ratings of each angry, happy, and neutral selected token. We ended up with a set of 90 tokens composed of 30 pseudowords recorded in each of the three emotional tones (anger, happiness, and neutrality).

Finally, 34 native Italian-speaking university students (17 men, 17 women, M age = 22.26, SD = 1.93, range: 19–26 years) rated each of the final 90 tokens for valence and arousal. After the audio presentation of a token, participants indicated their evaluation of the pronunciation on 21-point scales that ranged from −10 (*do not like it at all*) to +10 (*do like it very much*), and then specified the extent to which the pronunciation activated them emotionally on 11-point scales that ranged from 0 (*calm*) to 10 (*aroused/emotional*). This procedure was repeated for each token, presented in a random order. The interrater reliabilities for valence and arousal were good, αs = .92 and .91, respectively.

Emotional category and emotional perception, valence, and arousal of the selected 90 tokens

We conducted a series of 3 (emotion category) × 2 (speaker's gender) ANOVAs on the emotion, valence, and arousal ratings to evaluate the presence of significant differences in the emotional tone of the tokens (neutral, happy, angry with Bonferroni's corrected post-hoc tests) and of the gender of the speaker. In cases of significant interaction effects, we performed post-hoc analyses. We report all statistics in Table 2.

Emotional tone effects The angry, happy, and neutral tokens, on average, were perceived as having angrier, happier, and more neutral pronunciations, respectively. Moreover, the pronunciations of the angry tokens were evaluated as being the most negative, the pronunciations of the happy tokens were evaluated as being the most positive, and both of them were judged as being more arousing than the neutral tokens' pronunciations. Please refer to the Appendix (Tables 4 and 8) for the mean emotion, valence, and arousal ratings for each angry, happy, and neutral token of the selected pseudowords, pronounced by both the male and the female speaker.

Speaker's gender and Emotional Tone × Speaker's Gender interaction effects In terms of emotion perception, the speaker's gender affected the anger and neutrality ratings. Across tokens, the female pronunciation was judged as being more neutral and less angry than the male pronunciation. We also observed an Emotional Tone × Speaker's Gender interaction effect on the neutrality and arousal

Table 2 Emotion, valence, and arousal ratings and acoustic indexes of each emotional category of the tokens (overall and pronounced by a man or a woman)

	Happy Tokens			Angry Tokens			Neutral Tokens			Emotion		Gender		Interaction	
	Man M (SD)	Woman M (SD)	Total M (SD)	Man M (SD)	Woman M (SD)	Total M (SD)	Man M (SD)	Woman M (SD)	Total M (SD)	$F_{(1,84)}$	η_p^2	$F_{(1,84)}$	η_p^2	$F_{(2,84)}$	η_p^2
Anger ratings	1.68 (0.55)	1.52 (0.55)	1.60 (0.55)	16.96 (0.93)	16.53 (0.70)	16.74 (0.86)	5.59 (1.02)	4.73 (0.49)	5.16 (0.90)	3,376.98***	.99	9.41**	.10	1.69 ns	.04
Happiness ratings	14.87 (1.14)	14.72 (0.96)	14.80 (1.04)	1.55 (0.46)	1.73 (0.51)	1.64 (0.49)	7.04 (0.67)	6.38 (0.37)	6.71 (0.63)	2,413.31***	.98	1.79 ns	.02	2.41 ns	.05
Neutrality ratings	3.09 (1.27)	3.46 (0.88)	3.28 (1.09)	1.42 (0.51)	1.28 (0.35)	1.35 (0.44)	14.82 (1.34)	18.06 (.39)	16.44 (1.91)	2,563.87***	.98	38.11***	.31	31.54***	.43
Valence ratings	2.28 (0.62)	2.16 (1.09)	2.22 (0.88)	−3.16 (0.95)	−2.79 (0.70)	−2.98 (0.84)	0.46 (0.58)	0.27 (0.37)	0.36 (0.49)	361.28***	.90	0.02 ns	.00	1.23 ns	.03
Arousal ratings	5.53 (0.59)	5.54 (0.53)	5.53 (0.55)	5.63 (0.34)	5.74 (0.48)	5.69 (0.40)	3.70 (0.22)	3.18 (0.24)	3.44 (0.35)	270.84***	.87	2.12 ns	.03	5.16**	.11
f0	198.07 (31.93)	259.87 (24.53)	228.97 (42.07)	125.99 (24.29)	241.33 (33.30)	183.66 (65.27)	122.53 (13.97)	160.87 (8.58)	141.70 (25.58)	85.56***	.70	194.10***	.70	19.54***	.32
Intensity	82.57 (1.38)	83 (0.95)	82.79 (1.19)	80.23 (2.95)	83.49 (2.19)	81.86 (3.04)	80.10 (1.86)	80.15 (1.37)	80.13 (1.29)	16.64***	.28	10.71**	.11	7.03***	.14
Duration	0.88 (0.12)	0.64 (0.06)	0.76 (0.15)	1.21 (0.30)	0.83 (0.08)	1.02 (0.29)	0.97 (0.13)	0.81 (0.10)	0.89 (0.14)	21.41***	.34	67.26***	.44	4.03*	.09
f0 semitones	17.27 (4.58)	12.35 (4.01)	14.81 (4.91)	13.13 (3.64)	11.91 (2.49)	12.52 (3.13)	12.95 (3.78)	7.60 (1.43)	10.28 (3.91)	12.69***	.23	27.17***	.24	3.17*	.07

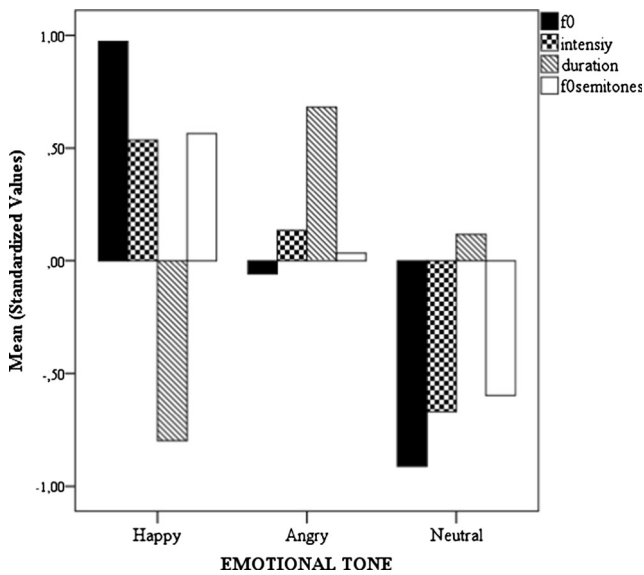* $p < .05$. ** $p < .01$. *** $p < .001$.

Fig. 1 Acoustic profiles of the three emotional categories of pseudowords

judgments. The neutrality judgments were higher and the arousal judgments were lower for neutral tokens than for angry or happy ones, but in a stronger way for those pronounced by a woman than for those pronounced by a man, suggesting that the male speaker might have had a less neutral intonation than the female speaker. Put together, these results can be interpreted in light of the literature about gender differences in the perception and expression of emotions. The literature has shown that sadness and fear are both considered stereotypical female emotions, whereas anger is primarily associated with masculinity (Fabes & Martin, 1991). In line with this theme, in a study on emotion expression and perception by men and women, Bonebright and colleagues (1996) found that male actors were perceived to portray anger better than female actors do.

Emotional category and acoustic properties of the selected 90 tokens

Similar to our analyses in the previous section, we conducted a series of 3 (emotion category) × 2 (speaker's gender) ANOVAs on each prosodic index with post-hoc analyses in cases of significant interaction effects. We report all statistics in Table 2.

Emotional tone effects Main effects of emotional tone were present on all of the prosodic indexes. The three emotional sets of tokens were all significantly different in terms of $f0$ and $f0$ semitones, with happy tokens showing the highest values, followed by angry and then neutral ones. Moreover, the happy and angry tokens showed significantly higher intensities than did neutral ones. Finally,

Table 3 Valences of the 40 preselected pseudowords (1 = *very negative*, 9 = *very positive*)

Pseudoword	M	SD	t(29)	p
ambato	5.03	2.14	0.09	.933
andori	4.77	1.72	−0.75	.462
badite	5.03	2.17	0.08	.934
balite	5.17	2.07	0.44	.662
balzeni	5.10	2.02	0.27	.789
blarela	4.67	2.35	−0.78	.444
elvuto	4.77	2.46	−0.52	.607
ervetto	5.40	2.43	0.90	.375
faddino	4.87	1.94	−0.38	.710
farani	4.90	1.83	−0.30	.766
faroli	5.00	1.70	0	1
fitasa	4.93	1.95	−0.19	.852
gaduri	4.90	2.06	−0.27	.792
gentormi	4.83	1.93	−0.47	.640
legeli	5.40	2.11	1.04	.308
lesersa	5.00	2.17	0	1
masteri	5.37	1.97	1.02	.317
mepaso	5.23	1.74	0.74	.467
montole	5.20	1.99	0.55	.586
ombrisso	4.83	1.95	−0.47	.643
ovilda	5.30	2.47	0.67	.510
padota	4.90	2.29	−0.24	.813
pavuma	4.93	1.96	−0.19	.854
pivrele	4.80	2.33	−0.47	.641
pramalo	4.73	1.98	−0.74	.467
psilumbo	5.33	2.60	0.70	.489
rantaglia	4.60	2.58	−0.85	.403
rollipi	4.83	2.26	−0.40	.689
ronape	4.60	2.06	−1.06	.297
sefrenta	4.57	2.28	−1.04	.307
tabempo	4.73	1.95	−0.75	.459
tambadi	4.87	1.81	−0.40	.690
tarandi	4.97	1.97	−0.09	.927
unvato	4.87	2.16	−0.34	.738
uvipa	4.97	1.99	−0.09	.928
vesita	5.07	1.78	0.20	.839
zellani	4.70	2.32	−0.71	.485
zimoma	5.37	2.08	0.97	.341
zompana	4.87	2.08	−0.35	.728
zorfara	4.60	2.16	−1.02	.319

the angry tokens had the longest durations, followed by the neutral ones, and then the happy ones. We illustrate the acoustic profiles in Fig. 1. Because of the different scales, we standardized all four acoustic measures. Overall, the acoustic profiles of our emotional sets of

pseudowords replicate previous findings (Banse & Scherer, 1996; Castro & Lima, 2010; Juslin & Laukka, 2003; Liu & Pell, 2012) in which happy tokens are characterized by the highest *f*0 values, high intensity, high pitch variability, and a faster speech rate, and angry tokens share pitch-level and intensity-level characteristics with happy tokens. However, in our set, angry tokens showed relatively less pitch variability and a lower speech rate, which contrasts with the typical anger acoustic profile described in the literature (Juslin & Laukka, 2003). This last result is in line with our analyses of the perceived emotion intensity and acoustic properties of all 270 tokens, and corresponds to Banse and Scherer's description of cold anger, as do the previously discussed results related to *f*0 variability. These results give further support to the hypothesis that our actors primarily conveyed cold anger in their prosodic expressions.

Speaker gender and Emotional Tone × Speaker's Gender interaction effects (see Table 2). We found speaker's gender effects on all acoustic properties, and all were qualified by an interaction with the emotional tone of the tokens. The tokens pronounced by the man had smaller *f*0, weaker intensities, longer durations, and higher semitones. The interaction effects evidenced different patterns of differences between the emotional tones, depending on the speaker. First, the higher *f*0 for neutral than for angry tokens was only observed for the female and not for the male speaker. Second, for the male speaker, happy tokens had a higher intensity than did angry ones, which, in turn, were not different from the neutral ones. However, for the female speaker, happy tokens had equal intensity with the angry ones, but both were higher than the neutral tokens' intensity. Third, for the male speaker, the duration of the angry tokens was longer than that of the happy and neutral ones, whereas for the female speaker, their duration was equal to that of the neutral tokens but longer than that of happy tokens. Finally, for the male speaker, angry tokens had lower semitones than did happy tokens, but the angry semitones were equal to the neutral ones. However, for the female speaker, the semitones of the angry tokens were not different from the happy semitones, but both were higher than the neutral semitones. Although some of these differences echo the natural variations in intonation among men and women (e.g., *f*0 differences; Childers & Wu, 1991; Puts, Doll, & Hill, 2014), others could be due to the idiosyncratic differences between the two actors, even though their voice types were quite prototypical in terms of *f*0.

Conclusion

The main goal of this study was to develop and validate a set of vocal emotional stimuli for research on emotional prosody processing following specific methodological criteria. First, we sought to generate pseudowords according to the specific phonotactic and distributional constraints of the Italian language. From this perspective, we relied on the Italian submodule of the Wuggy software (Keuleers & Brysbaert, 2010). The use of a formal validated criterion for pseudowords generation represents progress with respect to the available literature. The subjective judgment typically used in previous studies strongly depends on the researcher's experience with the specific languages and with pseudowords research. Another aim consisted of selecting only the neutral stimuli from independent valence ratings. The list of the selected pseudowords is available in the Appendix, Table 3. These stimuli could be very useful in elaborating lexical decision tasks, in which one needs to be sure that pseudowords are free of lexical–semantic and valence meanings. Finally, analyses based on independent ratings and the comparison between the acoustic properties of our dataset with well-established literature guarantee such stimuli specifically express the emotions they are intended to convey.

Although this set of stimuli would constitute ready-to-use and controlled material for experimental studies, we would like to note a series of methodological points. First, because we did not give specific cues to the actors in terms of how to express each emotion, it seems that they primarily focused on conveying cold anger. As was noted by Banse and Scherer (1996), most studies on prosodic emotion do not consider the difference between hot and cold anger, which could explain some discrepancies. On the basis of acoustic analyses, our set of angry pseudowords depicts cold and not hot anger. We would also underline the fact that our material was elaborated using only two actors, one man and one woman. Therefore, we cannot disentangle gender differences from idiosyncratic ones in some acoustic parameters of emotion expression that emerged in our results. However, this methodological issue should not prevent the use of our stimuli set, for two main reasons. The actors' voice types were prototypical of their gender, and the acoustic profiles of the selected set of stimuli they produced corresponded to what it is usually observed for male and female voices.

To conclude, our set of prosodic emotion expressions addresses several methodological issues that have affected previous databases. Furthermore, it overcomes the lack of validated stimuli for the Italian language, and it is available to the research community for future use in research involving emotional prosody from different perspectives (behavioral, clinical, neuropsychological), for both Italian samples and cross-language studies.

Appendix

Table 4 Anger, happiness, neutrality, valence, and arousal ratings of the Angry token of each of the final 15 pseudowords pronounced by a man

	Ratings			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
andori	16.85 (2.76)[a]	1.28 (1.94)[b]	1.50 (2.47)[bc]	5.05 (2.66)	−3.32 (4.03)
balite	16.43 (3.33)[a]	1.39 (1.90)[b]	1.59 (1.95)[bc]	5.44 (2.24)	−1.91 (5.01)
blarela	17.78 (2.51)[a]	1.35 (2.50)[b]	0.82 (1.18)[bc]	5.76 (2.48)	−3.55 (5.52)
farani	18.35 (1.85)[a]	0.52 (1.05)[b]	0.74 (1.65)[bc]	6.35 (2.38)	−4.67 (4.39)
fitasa	16.59 (3.77)[a]	2 (2.99)[b]	1.39 (1.83)[bc]	5.20 (2.78)	−3.91 (4.80)
gentormi	18.15 (2.23)[a]	1.46 (2.69)[b]	0.61 (1.27)[bc]	5.79 (2.90)	−4.58 (4.76)
lesersa	17.04 (3.10)[a]	1.35 (2.52)[b]	1.35 (2.01)[bc]	5.5 (2.45)	−2.26 (5.63)
rantaglia	17.78 (2.15)[a]	1.75 (4.63)[b]	1.44 (3.76)[bc]	5.97 (2.22)	−3.05 (4.95)
rollipi	16.48 (2.79)[a]	1.39 (2.27)[b]	1.33 (2.34)[bc]	5.61 (2.26)	−2.52 (5.42)
sefrenta	15.37 (3.59)[a]	2.24 (2.29)[b]	2.07 (3.04)[bc]	5.73 (2.33)	−1.47 (4.84)
unvato	15.96 (2.69)[a]	1.67 (2.06)[b]	2.07 (2.25)[bc]	5.35 (2.29)	−2.55 (4.64)
vesita	17.87 (2.82)[a]	1.09 (2.18)[b]	0.76 (1.95)[bc]	5.61 (2.64)	−2.94 (5.96)
zellani	17.41 (4.28)[a]	2.09 (4.97)[b]	1.53 (3.80)[bc]	5.88 (2.30)	−3.91 (5.91)
zimoma	15.48 (3.72)[a]	2.17 (2.41)[b]	2.20 (3.23)[bc]	5.35 (2.17)	−2.70 (4.75)
zorfara	16.85 (4.02)[a]	1.50 (2.35)[b]	1.87 (3.51)[bc]	5.73 (2.27)	−4.02 (5.64)

Means not sharing a letter within ratings (angry, happy, and neutral) are significantly different ($p < .05$).

Table 5 Anger, happiness, neutrality, valence, and arousal ratings of the happy token of each of the final 15 pseudowords pronounced by a man

	Ratings			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
andori	0.80 (1.69)[a]	14.83 (3.38)[b]	2.99 (3.26)[c]	4.97 (2.81)	3.11 (4.70)
balite	2.28 (3.59)[a]	14.24 (3.88)[b]	2.65 (3.30)[ac]	5.82 (2.02)	1.5 (4.50)
blarela	2.07 (3.43)[a]	15.41 (3.92)[b]	2.63 (3.69)[ac]	5.94 (2.37)	2.17 (5.32)
farani	2.98 (3.85)[a]	12.26 (4.19)[b]	4.57 (4.57)[ac]	5.23 (2.25)	1.85 (3.20)
fitasa	1.70 (3.22)[a]	16.35 (3.37)[b]	1.89 (3.05)[ac]	5.97 (1.99)	2.64 (4.55)
gentormi	1.11 (2.11)[a]	15.30 (2.97)[b]	3.20 (3.17)[c]	5.20 (2.37)	2.14 (3.93)
lesersa	1.48 (2.22)[a]	13.61 (4.07)[b]	5.65 (4.48)[c]	4.88 (2.21)	2 (3.85)
rantaglia	1.94 (3.51)[a]	14.63 (3.59)[b]	3.50 (3.30)[c]	4.85 (2.13)	2.05 (4.26)
rollipi	1.67 (3.67)[a]	15.50 (3.83)[b]	1.63 (2.76)[ac]	6.08 (2.46)	2.94 (3.79)
sefrenta	0.89 (1.96)[a]	16.22 (2.92)[b]	1.83 (2.96)[ac]	6 (2.60)	2.97 (4.68)
unvato	1.35 (2.24)[a]	15.11 (3.73)[b]	3.43 (3.81)[c]	5.85 (2.00)	1.88 (3.70)
vesita	1.50 (2.51)[a]	13.57 (2.94)[b]	4.93 (4.37)[c]	4.52 (2.14)	1.32 (3.80)
zellani	1.94 (3.88)[a]	14.78 (3.14)[b]	3.69 (3.75)[ac]	5.26 (2.31)	2.35 (4.14)
zimoma	1.78 (3.41)[a]	14.72 (4.14)[b]	2.83 (3.62)[ac]	5.55 (2.71)	1.76 (4.40)
zorfara	1.76 (3.97)[a]	16.54 (4.49)[b]	1.02 (2.30)[ac]	6.70 (2.16)	3.47 (4.66)

Means not sharing a letter within emotion ratings (angry, happy, and neutral) are significantly different ($p < .05$).

Table 6 Anger, happiness, neutrality, valence, and arousal ratings of the neutral token of each of the final 15 pseudowords pronounced by a man

	Ratings			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
andori	5.83 (4.98)[a]	5.72 (3.78)[ab]	16.07 (3.95)[c]	3.38 (2.36)	0.29 (3.20)
balite	5.91 (4.98)[a]	6.41 (3.54)[ab]	13.59 (4.92)[c]	3.85 (2.47)	−0.7 (3.50)
blarela	7.80 (5.36)[a]	6.22 (3.62)[ab]	12.76 (5.26)[c]	3.88 (2.12)	0.20 (3.62)
farani	4.59 (4.47)[a]	7.80 (4.11)[b]	15.98 (3.91)[c]	3.5 (2.47)	0.44 (3.03)
fitasa	5.87 (4.43)[a]	7.48 (3.98)[ab]	13.70 (4.96)[c]	3.52 (2.24)	0.94 (3.45)
gentormi	6.22 (4.34)[a]	6.80 (3.80)[ab]	14.17 (5.25)[c]	3.44 (2.53)	0.76 (3.55)
lesersa	4.08 (4.40)[a]	7.63 (3.23)[b]	15.54 (4.59)[c]	4 (2.25)	1.41 (3.61)
rantaglia	4.75 (4.85)[a]	6.22 (4.29)[ab]	16.03 (4.67)[c]	3.70 (2.18)	0.35 (3.90)
rollipi	5.89 (4.45)[a]	7.72 (3.75)[ab]	13.96 (5.44)[c]	3.91 (2.42)	−0.4 (3.81)
sefrenta	5.07 (4.54)[a]	7.15 (3.54)[b]	15.61 (4.19)[c]	3.67 (2.19)	0.20 (3.43)
unvato	6.83 (5.29)[a]	7.50 (3.43)[ab]	13.65 (5.02)[c]	3.70 (2.16)	0.44 (2.77)
vesita	6.20 (5.09)[a]	7.70 (3.44)[ab]	12.93 (5.17)[c]	3.64 (2.26)	0.91 (2.63)
zellani	5.41 (4.42)[a]	6.56 (4.18)[ab]	15.09 (4.11)[c]	3.58 (2.24)	0 (3.32)
zimoma	3.89 (4.12)[a]	7.61 (3.77)[b]	16.87 (4.68)[c]	3.55 (2.41)	1.17 (3.16)
zorfara	5.59 (4.43)[a]	7.04 (3.75)[ab]	16.33 (4.16)[c]	4.14 (2.04)	0.91 (2.90)

Means not sharing a letter within emotion ratings (angry, happy, and neutral) are significantly different ($p < .05$).

Table 7 Anger, happiness, neutrality, valence, and arousal ratings of the angry token of each of the final 15 pseudowords pronounced by a woman

	Emotion			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
ambato	16.50 (3.87)[a]	1.74 (3.38)[b]	0.89 (1.66)[bc]	6.47 (2.17)	−1.94 (5.54)
elvuto	16.46 (3.20)[a]	1.61 (2.02)[b]	1.20 (2)[bc]	5.85 (2.54)	−2.70 (4.95)
faddino	15.70 (3.12)[a]	1.85 (2.14)[b]	1.67 (2.18)[bc]	5.73 (2.01)	−3.85 (4.25)
gaduri	15.80 (3.09)[a]	1.52 (1.86)[b]	1.67 (3.06)[bc]	5.38 (2.44)	−3.70 (4.05)
legeli	15.89 (3.45)[a]	1.80 (2.31)[b]	1.70 (2.63)[bc]	5.52 (2.12)	−2.58 (4.34)
masteri	16.57 (2.54)[a]	1.20 (1.57)[b]	1.07 (1.79)[bc]	6.02 (2.46)	−2.08 (4.22)
montole	17.72 (2.33)[a]	1.24 (2.37)[b]	1.02 (1.81)[bc]	6.26 (2.15)	−1.91 (5.11)
ovilda	16.11 (3.40)[a]	2.07 (2.30)[b]	1.30 (1.72)[c]	4.97 (2.22)	−2.61 (3.91)
pavuma	15.98 (2.35)[a]	2.04 (2.38)[b]	1.43 (1.93)[bc]	5.32 (2.25)	−2.20 (4.65)
pramalo	17.20 (2.33)[a]	1.72 (2.41)[b]	1.22 (1.75)[bc]	5.73 (2.32)	−3.05 (4.27)
ronape	18.02 (2.70)[a]	1.17 (3.01)[b]	0.54 (1.15)[bc]	6.11 (2.34)	−3.64 (4.79)
tabempo	16.46 (2.83)[a]	1.28 (2.02)[b]	1.37 (2.02)[bc]	5.44 (2.53)	−3.64 (4.12)
tarandi	16.85 (2.84)[a]	1.93 (2.55)[b]	1.43 (2.44)[bc]	5.94 (2.11)	−3.35 (4.24)
uvipa	15.43 (2.50)[a]	3.24 (3.46)[b]	1.78 (2.36)[c]	5.11 (2.14)	−2.11 (3.70)
zompana	17.28 (2.34)[a]	1.50 (3.11)[b]	0.91 (1.44)[bc]	6.29 (2.46)	−2.38 (5.17)

Means not sharing a letter within ratings (angry, happy, and neutral) are significantly different ($p < .05$).

Table 8 Anger, happiness, neutrality, valence, and arousal ratings of the happy token of each of the final 15 pseudowords pronounced by a woman

	Ratings			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
ambato	1.33 (2.14)[a]	14.63 (3.85)[b]	3.50 (3.84)[c]	4.82 (2.79)	0.88 (5.19)
elvuto	1.59 (2.90)[a]	14.07 (3.45)[b]	3.15 (3.50)[c]	5.64 (2.11)	2.17 (4.24)
faddino	1.87 (3.36)[a]	13.78 (3.35)[b]	3.89 (3.74)[c]	5.58 (1.84)	1.23 (4.17)
gaduri	2 (3.30)[a]	14.28 (3.10)[b]	4.70 (4.38)[c]	5.23 (2.20)	1.67 (3.32)
legeli	1.67 (2.70)[a]	14.83 (3.35)[b]	3.28 (3.96)[c]	5.5 (2.09)	1.85 (4.25)
masteri	1.48 (2.49)[a]	14.46 (3.54)[b]	3.72 (3.59)[c]	5.58 (2.07)	2.17 (3.40)
montole	1.07 (2.30)[a]	15.98 (2.88)[b]	2.54 (3.37)[c]	6.32 (1.96)	3.91 (4.37)
ovilda	1.02 (1.86)[a]	15.02 (2.71)[b]	2.98 (3.67)[c]	5.91 (2.19)	3.20 (4.17)
pavuma	2.09 (2.66)[a]	12.41 (3.51)[b]	5.67 (4.89)[c]	4.5 (1.71)	0.64 (3.67)
pramalo	1.80 (2.66)[a]	14.04 (3.61)[b]	4.15 (4.28)[c]	4.70 (2.15)	0.32 (3.51)
ronape	0.80 (2.10)[a]	14.96 (3.66)[b]	2.74 (3.40)[c]	5.58 (2.24)	2.70 (3.50)
tabempo	0.87 (2.20)[a]	15.91 (3.06)[b]	2.39 (2.98)[c]	5.58 (2.51)	3.79 (4.11)
tarandi	2.87 (4.03)[a]	15.09 (3.12)[b]	3.61 (4.07)[ac]	6.14 (1.97)	2.08 (3.89)
uvipa	1.24 (2.94)[a]	16.07 (2.79)[b]	2.80 (3.29)[c]	5.97 (2.62)	3.11 (3.46)
zompana	1.17 (2.51)[a]	15.30 (2.97)[b]	2.76 (3.46)[c]	6.02 (2.34)	2.61 (4.22)

Means not sharing a letter within ratings (angry, happy, and neutral) are significantly different ($p < .05$).

Table 9 Anger, happiness, neutrality, valence, and arousal ratings of the neutral token of each of the final 15 pseudowords pronounced by a woman

	Ratings			Arousal	Valence
	Anger M (SD)	Happiness M (SD)	Neutrality M (SD)	M (SD)	M (SD)
ambato	4.50 (4.45)[a]	6.26 (3.65)[ab]	18.41 (2.86)[c]	2.64 (2.46)	0.70 (2.73)
elvuto	5.35 (4.87)[a]	6.74 (3.45)[ab]	17.98 (3.43)[c]	3.52 (2.41)	0.29 (3.24)
faddino	4.52 (4.14)[a]	6.91 (3.51)[b]	17.72 (3.82)[c]	3.26 (2.31)	0.11 (2.44)
gaduri	4.78 (4.09)[a]	6.48 (3.74)[ab]	17.87 (3.40)[c]	3.05 (2.39)	−0.20 (2.67)
legeli	3.87 (3.95)[a]	7.09 (3.98)[b]	18.21 (3.18)[c]	2.91 (2.28)	0.67 (2.99)
masteri	4.30 (4.31)[a]	6.41 (3.54)[b]	17.83 (3.73)[c]	3.02 (2.27)	0.26 (2.90)
montole	5.35 (4.50)[a]	6.02 (3.65)[ab]	17.41 (4.85)[c]	3.11 (2.18)	0.76 (2.60)
ovilda	4.11 (4.68)[a]	6.37 (3.69)[b]	18.43 (2.97)[c]	3.14 (2.42)	0.58 (2.61)
pavuma	5.09 (4.58)[a]	6.17 (3.60)[ab]	18.20 (3.36)[c]	3.23 (2.29)	0.20 (3.47)
pramalo	4.93 (4.64)[a]	6.43 (3.72)[ab]	17.93 (3.36)[c]	3.35 (2.33)	0 .23 (3.54)
ronape	4.70 (4.56)[a]	6.65 (3.78)[b]	18.13 (3.15)[c]	3.61 (2.44)	0 (3.11)
tabempo	4.28 (4.06)[a]	5.85 (3.94)[ab]	17.46 (4.62)[c]	2.97 (2.46)	0 (3.03)
tarandi	4.54 (4.36)[a]	5.76 (3.65)[ab]	18.22 (2.96)[c]	3.32 (2.27)	0.82 (3.64)
uvipa	5.13 (4.54)[a]	6.43 (3.70)[ab]	18.91 (2.23)[c]	3.29 (2.34)	−0.30 (2.93)
zompana	5.54 (4.72)[a]	6.15 (3.68)[ab]	18.15 (3.67)[c]	3.17 (2.36)	0.11 (2.39)

Means not sharing a letter within ratings (angry, happy, and neutral) are significantly different ($p < .05$).

References

Ariatti, A., Benuzzi, F., & Nichelli, P. (2008). Recognition of emotions from visual and prosodic cues in Parkinson's disease. *Neurological Science, 29,* 219–227. doi:10.1007/s10072-008-0971-9

Bach, D. R., Grandjean, D., Sander, D., Herdener, M., Strik, W. K., & Seifritz, E. (2008). The effect of appraisal level on processing of emotional prosody in meaningless speech. *NeuroImage, 42,* 919–927. doi:10.1016/j.neuroimage.2008.05.034

Bachorowski, J. A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science, 6,* 219–224.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70,* 614–636.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5,* 341–345.

Bonebright, T. L., Thompson, J. L., & Leger, D. W. (1996). Gender stereotypes in the expression and perception of vocal affect. *Sex Roles, 34,* 429–445.

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., & Weiss, B. (2005). A database of German emotional speech. In *Proceedings of the 9th European Conference on Speech Communication and Technology (Interspeech 2005)* (pp. 1517–1520). Bonn, Germany: International Speech Communication Association.

Castagna, F., Montemagni, C., Milani, A. M., Rocca, G., Rocca, P., Casacchia, M., & Bogetto, F. (2013). Prosody recognition and audiovisual emotion matching in schizophrenia: The contribution of cognition and psychopathology. *Psychiatry Research, 205,* 192–198. doi:10.1016/j.psychres.2012.08.038

Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods, 42,* 74–81. doi:10.3758/BRM.42.1.74

Childers, G., & Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis. *Journal of the Acoustical Society of America, 90,* 1841–1856.

Dellacherie, D., Hasboun, D., Baulac, M., Belin, P., & Samson, S. (2011). Impaired recognition of fear in voices and reduced anxiety after unilateral temporal lobe resection. *Neuropsychologia, 49,* 618–629. doi:10.1016/j.neuropsychologia.2010.11.008

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1,* 56–75.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *The Facial Action Coding System* (2nd ed.). London, UK: Weidenfeld & Nicolson.

Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin, 17,* 532–540.

Hagenhoff, M., Franzen, N., Gerstner, L., Koppe, G., Sammer, G., Netter, P., & Lis, S. (2013). Reduced sensitivity to emotional facial expressions in borderline personality disorder: Effects of emotional valence and intensity. *Journal of Personality Disorders, 27,* 19–35. doi:10.1521/pedi.2013.27.1.19

Hawk, S. T., van Kleef, G. A., Fischer, A. H., & van der Schalk, J. (2009). "Worth a thousand words": Absolute and relative decoding of non-linguistic affect vocalizations. *Emotion, 9,* 293–305. doi:10.1037/a0015178

Henton, C. G. (1989). Fact and fiction in the description of female and male pitch. *Language & Communication, 9,* 299–311.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage, 58,* 665–674. doi:10.1016/j.neuroimage.2011.06.035

Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happé, F., Scott, S. K., & Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 52*(3), 275-285.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion, 1,* 381–412. doi:10.1037/1528-3542.1.4.381

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*, 770–814. doi:0.1037/0033-2909.129.5.770

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42,* 627–633. doi:10.3758/BRM.42.3.627

Kirsh, S. J., & Mounts, J. R. (2007). Violent video game play impacts facial emotion recognition. *Aggressive Behavior, 33,* 353–358.

Kotz, S. A., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research, 1151,* 107–118. doi:10.1016/j.brainres.2007.03.015

Kotz, S., Meyer, M., Alter, K., Besson, M., von Cramon, D., & Friederici, A. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language, 86,* 366–376.

Liu, P., & Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behavior Research Methods, 44,* 1042–1051. doi:10.3758/s13428-012-0203-3

Minzenberg, M. J., Poole, J. H., & Vinogradov, S. (2006). Social-emotion recognition in borderline personality disorder. *Comprehensive Psychiatry, 47,* 468–474. doi:10.1016/j.comppsych.2006.03.005

Mitchell, R. L., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41,* 1410–1421.

Paulmann, S., Pell, D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: Evidence from ERPs. *Brain Research, 1217,* 171–178. doi:10.1016/j.brainres.2008.04.032

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience, 30,* 10127–10134. doi:10.1523/JNEUROSCI.2161-10.2010

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *Journal of the Acoustical Society of America, 109,* 1668–1680.

Pell, M. D. (2002). Evaluation of nonverbal emotion in face and voice: Some preliminary findings on a new battery of tests. *Brain and Cognition, 48,* 499–504.

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics, 37,* 417–435. doi:10.1016/j.wocn.2009.07.005

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social cognition in borderline personality disorder: Evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience, 4,* 182. doi:10.3389/fnbeh.2010.00182

Puts, D. A., Doll, L. M., & Hill, A. K. (2014). Sexual selection on human voices. In V. Weekes-Shackelford & T. K. Shackelford (Eds.), *Evolutionary perspectives on human sexual psychology and behavior* (pp. 69–86). Heidelberg, Germany: Springer. doi:10.1007/978-1-4939-0314-6_3

Schaffer, S. G., Wisniewski, A., Dahdah, M., & Froming, K. B. (2009). The comprehensive affect testing system-abbreviated: Effects of age on performance. *Archives of Clinical Neuropsychology, 24,* 89–104. doi:10.1093/arclin/acp012

Scherer, K. R. (1989). *Vocal correlates of emotional arousal and affective disturbance*. In H. Wagner & A. Manstead (Eds.), Handbook of social psychophysiology (pp. 165–197). New York: Wiley.

Scherer, K. R., & Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the emotion recognition index. *Journal of Nonverbal Behavior, 4*, 305–326. doi:10.1007/s10919-011-0115-4

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14,* 228–233.