Affective prosody in European Portuguese: Perceptual and acoustic characterization of one-word utterances

Marisa G. Filipe [a,*], Paulo Branco [a], Sónia Frota [b], São Luís Castro [a,*], Selene G. Vicente [a,*]

[a] *Speech Lab, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal*
[b] *Center of Linguistics, University of Lisbon, Lisbon, Portugal*

Abstract

A perceptual and acoustic characterization was provided on the expression of liking and disliking in the European Portuguese language. Thirty participants identified vocal patterns and judged the intensity of expressed affect in one-word items recorded by six untrained speakers. Listeners consistently associated vocal profiles with the two emotional patterns of liking and disliking. However, liking intonation was easier to recognize than disliking intonation. The feature most commonly associated with liking intonation was a wider and higher F0 pattern and a rising-falling contour. For disliking, the results revealed a flatter melodic pattern with a fall into the stressed syllable yielding a low plateau. In sum, both prosodic patterns showed different and consistent correlates.
© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Emotional prosody refers to the variation in acoustic cues such as fundamental frequency (F0), amplitude (or intensity), timing, and voice quality during speech that is used to convey the emotional meaning of an utterance (e.g., Kotz and Paulmann, 2011). The role of vocal cues in the expression of feelings and affection, as well as their powerful effects on social interaction, has been recognized for many years (e.g., Darwin, 1872/1965). Indeed, beginning with the first moments of life, emotional prosody is commonly used by developing individuals to express and perceive emotions. This ability is crucial to communication, and it might be atypical or impaired in some clinical cases, such as in children with speech and language impairments (Wells and Peppé, 2003). Consequently, it is both socially

and clinically relevant to understand how individuals realize, for instance, that someone is irritated, frightened, or pleased. Because the vocal expression of affection and emotions is associated with perceptual identification and acoustic profiles (for a review, see Juslin and Laukka, 2003), the present study examined perceptual and acoustic features involved in the vocal expression of liking and disliking in the European Portuguese language.

Although the communication of emotions through facial expression has received much attention, psychologists, linguists, psycholinguists, and neuroscientists have been increasingly focusing on the vocal communication of emotions (e.g., Elfenbein et al., 2002). There is currently a significant body of knowledge on emotional vocal cues (e.g., Juslin and Laukka, 2003; Scherer, 2003). Several main questions have already been addressed. First, several studies indicated that verbal comprehension is affected by prosody. For example, Nygaard and Lunders (2002) found that a listener's selection of word meaning in homophones (e.g., die/dye) with and without emotional meaning was influenced by the emotional tone of voice (see also: Ishii

* Corresponding authors at: Speech Lab, Faculty of Psychology and Education Sciences, University of Porto, Rua Alfredo Allen, 4200-135 Porto, Portugal. Tel.: +351 226 079 700; fax: +351 226 079 725.
E-mail address: labfala@fpce.up.pt (M.G. Filipe).

et al., 2003; Kitayama and Ishii, 2002). Neuroimaging studies have found general dissociations between emotional prosody and linguistic processing. For instance, individuals with damage to the right hemisphere have difficulty in identifying emotional prosody, although their ability to extract the literal linguistic content of the spoken utterance remains intact (Adolphs et al., 2002). Therefore, it is important to take into account both linguistic and prosodic content when determining the emotional perception of a word.

Another issue that has been addressed is the search for vocal profiles that are associated with particular emotions. Juslin and Laukka (2001, 2003), among others (e.g., Banse and Scherer, 1996; Hammerschmidt and Jürgens, 2007; Paulmann et al., 2008; Scherer et al., 1991), have shown that there are specific acoustic profiles associated with the vocal expression of different emotion categories. For instance, happiness is characterized by high, rising, and highly variable F0, fast speech rate, medium–high voice intensity, medium high-frequency energy, and fast voice onsets. In contrast, sadness can be characterized by low, falling, and not so variable F0, slow speech rate, low voice intensity, little high-frequency energy, and slow voice onsets/attacks. In fact, acoustic variables such as speech rate, voice intensity, voice quality, and F0 (mean, maximum, minimum, and deviation) can be powerful means to categorize emotions (Juslin and Madison, 1999), and acoustic tools (i.e., discriminant acoustic vocal features) have been proven efficient in profiling emotions and in assisting emotional recognition.

Research has also demonstrated that listeners reliably and accurately perceive emotions in the human voice. Typically, listeners select the emotion that best describes the stimulus from a list of emotion labels (forced-choice recognition task), and then the percentage of correctly recognized stimuli per emotion is calculated. These perceptual classifications suggest that listeners detect emotion in voice with great accuracy (Castro and Lima, 2010; Juslin and Laukka, 2003; Pell et al., 2009; Scherer, 2003; Scherer et al., 2003; Sobin and Alpert, 1999). As a result, most studies on emotional prosody used emotional identification tasks and asked the participants to identify the emotion presented in speech samples (e.g., Banse and Scherer, 1996; Scherer, 1979). For example, Castro and Lima (2010), with adult speakers of Portuguese, showed that discrete emotions (e.g., anger, disgust, fear, happiness, sadness, surprise, and neutrality) were recognized well-above chance and were given high-intensity ratings. However, not all emotions seem to be decoded equally well. Banse and Scherer (1996) found that anger and sadness were decoded more reliably than joy and fear. Considering the high levels of accuracy in the identification of most emotions based on vocal parameters, as well as findings suggesting that vocal profiles are associated with particular emotions, the assumption of distinct profiles for emotions based on acoustic and perceptual features is strongly supported in the literature.

Recently, research has focused on how the vocal expression of emotions varies among languages and cultures. Although some researchers have suggested that there are innate, stable, and universal features (e.g., Juslin and Laukka, 2003), differences have also been reported. For example, liking and disliking have been characterized differently across languages. For British English, this distinction can be expressed in the use of differing tones, rise-fall for liking and fall-rise for disliking (Martinez-Castilla and Peppé, 2008). Also, Banse and Scherer (1996) demonstrated that English utterances produced with a wider and higher pitch range suggest positive affect. On the other hand, for Spanish, Martinez-Castilla and Peppé (2008) found that the intonation feature most indicative of liking was an inverted U-shaped contour with an exaggerated pitch peak on the stressed syllable, whereas disliking was mostly expressed by a flat melodic contour with a slight fall on the stressed syllable. Thus, there may be differences between languages in the expression of affective valences. Research has not yet addressed this topic extensively, and issues about language differences remain unclear.

In sum, the vocal expression of affect is reliably associated with acoustic cues and elicits reliable listener identification, and researchers in the field have conducted acoustic and perceptual analyses to establish the vocal patterning of expressed emotions. However, important issues remain to be addressed. Further research is needed to answer the question of whether the recognition of affect differs across languages, and to what extent the vocal cues have some degree of language specificity. European Portuguese includes properties of both Romance and Germanic languages in its phonology and prosody (Frota, 2000, 2014; see also Duncan et al., 2013), posing challenging questions for the study of emotions, and no study so far has characterized liking and disliking vocal patterns for this language. The present study was designed to address the following issues: How well are liking and disliking recognized on the basis of prosody, and which intonation features are involved in the expression of liking and disliking in general, and specifically in European Portuguese (EP)?

2. Method

2.1. Participants

Six native speakers of EP (from the Porto region; three women and three men) between 22 and 47 years ($M = 31.83$, $SD = 10.22$) recorded the production data. None of these speakers had had any specific vocal training. Another 30 undergraduate students (24 women and 6 men) studying psychology at the University of Porto, with a mean age of 19.3 years ($SD = 0.95$) participated in the perceptual experiment. All were native speakers of EP with no hearing difficulties or speech disorders.

2.2. Materials

A set of six one-word items for training plus 22 one-word items for the acoustic and perceptual test (i.e., experimental condition) were selected, taking into account

criteria for high frequency and age of acquisition equal or inferior to 5 years (Cameirão and Vicente, 2010; Gomes and Castro, 2003; the complete list is in Appendix A). The words consisted of food-related items (e.g., *laranja*, orange) with different lengths in number of syllables (from 1 to 3). Most words had penultimate stress, since this is the most common stress pattern in Portuguese. For training, 2 words for each word length were used; for the experimental items, there were 11 disyllables, 10 three-syllable-words, and 1 monosyllable.

The 28 items were recorded by the six speakers. Six separate recording sessions of about one hour each were conducted, one for each speaker. All speakers were asked to produce the items in two different affective tones: liking and disliking. The instructions were, (A) "Say the name of the food item as if you like it" and (B) "Say the name of the food item as if you don't like it". Stimuli were recorded with a high-quality microphone in a soundproof room using the speech station of the Speech Laboratory at the University of Porto. The software used was Pro Tools LE version 5.1.1 with a 48-kHz sampling rate and 16-bit resolution.

Post-processing of the data included slicing the recording sessions into individual sound files containing one-word utterances, and setting onset and end of sound files to 100 ms silence. A total of 12 training items (6 liking + 6 disliking), two from each speaker, were randomly selected, and then 264 experimental items (22 items × 2 patterns × 6 speakers) were chosen for the perceptual test and for acoustic analysis. These items were divided into two blocks [Block A (132 items) + Block B (132 items)] with a 1-min pause between blocks and an inter-stimulus interval of 4 s. Item order was randomized.

2.3. Procedure

For the perceptual test, participants were individually tested in one experimental session that lasted about 20 min. The training block ($N = 12$) was presented first in order to familiarize the subject with the task and to avoid learning curve effects. It was followed by the two blocks of experimental trials. Reaction times and accuracy of the responses were recorded with SuperLab V4.0 (Abboud et al., 2006). All items were presented through high-quality headphones. Participants were asked to perform two independent, consecutive judgments for each stimulus: (1) classify each stimulus as neutral, expressing liking, or expressing disliking (forced-choice task) and (2) grade the perceived intensity of the expressed emotion on a seven-point scale from 1, *not intense,* to 7, *very intense.*

Additionally, an acoustic analysis of F0 (F0 mean, F0 range, F0 maximum, F0 minimum, and F0 contour), intensity, and duration was conducted for the 264 experimental items using PRAAT (Boersma and Weenink, 2011). The parameters were chosen because they have been widely investigated in the expression of affective patterns. Also, we adopted the Autosegmental-Metrical Model of intonational phonology (Ladd, 2008) and the Tones and Break Indices framework (Beckman et al., 2005), which have been used for several languages including Portuguese (Frota, 2014) in order to perform a phonological analysis of the nuclear contour patterns obtained in the data.

3. Results

3.1. Perceptual analysis

The initial analysis explored the accuracy in identifying liking and disliking patterns. Accuracy measured in percent of correct identifications was calculated for each speaker (see Table 1). Mean identification accuracy ranged from 73% to 91%, showing that participants were able to identify the affective patterns with great accuracy. Moreover, inter-participant consistency in the identification of the intended patterns was very high for every speaker: Cronbach's alpha was 0.92. Regarding the intensity of the ratings, the vast majority of the items were classified as having 5 or more points on the seven-point scale. There was a moderate Pearson correlation between the percentage of correct identifications and the intensity of expressed affect ($r = .66$; $p < .001$), showing that the higher the perceived intensity, the higher the accuracy in identification.

We tested the normality of all of the variables (accuracy, F0 mean, F0 range, maximum F0, minimum F0, duration, and intensity) with the Kolmogorov–Smirnov Test, and the data were not normally distributed. We then inspected skewness and kurtosis, which are common criteria for assessing normality, and the resulting values fell under 1 for all of the variables, well below the values recommended as acceptable by Kline (2005). This led us to use parametric methods.

One of our interests was to examine the accuracy of the participants' answers according to word length. Monosyllables had an overall accuracy of 79%, disyllables 82%, and 3-syllable words 85%. We found that word length had no influence on accuracy ($F (2, 251) = 1.83$; *ns.*). Regarding the affective patterns, liking items obtained about 92% of correct answers, and disliking items 75%. Analyses of variance (ANOVA) showed that liking items were significantly easier to recognize than disliking items ($F (1, 251) = 54.19$; $p < .001$; $\eta^2 = .178$). Finally, words produced by men (85% correct answers) were easier to

Table 1

Distribution (mean%) of the participants' responses ($N = 22$) for the total of the 264 items presented, separately for each speaker ($N = 6$).

Speakers	Gender	Responses		
		% Correct	% Incorrect	Mean intensity rating
1	Female	77.6	22.4	5.71
2	Female	72.6	27.4	5.38
3	Female	90.5	9.6	5.83
4	Male	88.5	11.5	5.92
5	Male	82.2	17.8	5.30
6	Male	84.2	15.8	5.61

Table 2
Mean (M) and standard deviation (SD) of the fundamental frequency (F0), F0 range (SD), Maximum F0, minimum F0, and duration.

Variables	Mean F0 (Hz)	F0 SD (Hz)	Max. F0 (Hz)	Min. F0 (Hz)	Duration (sec)
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Speaker					
Speaker 1	193 (11)	37 (7)	289 (26)	134 (17)	1.37 (0.19)
Speaker 2	190 (22)	31 (11)	257 (42)	133 (28)	1.23 (0.22)
Speaker 3	215 (16)	44 (11)	308 (16)	124 (20)	1.49 (0.17)
Speaker 4	166 (18)	37 (11)	261 (46)	118 (10)	1.47 (0.19)
Speaker 5	125 (18)	29 (12)	183 (40)	81 (4)	1.29 (0.23)
Speaker 6	164 (43)	36 (21)	225 (67)	108 (13)	1.35 (0.19)
Gender					
Women	199 (20)	37 (11)	284 (37)	130 (23)	1.36 (0.23)
Men	152 (35)	34 (15)	223 (61)	102 (18)	1.37 (0.22)
Emotion					
Like	190 (27)	41 (10)	280 (37)	122 (23)	1.34 (0.19)
Dislike	160 (39)	31 (15)	227 (64)	111 (26)	1.39 (0.24)
Word length					
Monosyllable	180 (41)	34 (16)	250 (68)	122 (25)	1.12 (0.19)
Disyllable	175 (37)	34 (14)	251 (61)	119 (25)	1.29 (1.17)
3-Syllable	174 (35)	37 (13)	257 (55)	113 (35)	1.50 (0.19)

recognize than the words produced by women (80% correct answers; F (1, 251) = 4.52; $p < .05$; $\eta^2 = .018$), although with an extremely small effect size. In sum, our results demonstrated that accuracy was high, and liking prosody was more easily recognized. Variability in the vocal cues and individual differences involved in the expression of liking and disliking were analyzed next.

3.2. Acoustic analysis

The mean and standard deviations of the acoustic measures – F0 mean, F0 range, maximum F0, minimum F0, and duration – are presented in Table 2. As expected, there were significant differences between women and men in all F0 measurements. To control for gender variance, we used z-transformed residuals of the acoustic parameter scores. As shown in Table 3, the z-scores independent of gender indicated a positive deviance towards the mean in the liking condition, whereas disliking had a negative deviance from the mean. Liking and disliking differed from one another in F0 mean, F0 range, F0 maximum, F0 minimum, and duration, as revealed by one-way ANOVAs and the robust test of equality of means with the Brown-Forsythe statistic (F (1, 251) = 50.16; $p < .001$; $\eta^2 = .161$; F (1, 251) = 36.56; $p < .001$; $\eta^2 = .123$; F (1, 251) = 66.38; $p < .001$; $\eta^2 = .202$; F (1, 251) = 11.66; $p = .001$; $\eta^2 = .043$; F (1, 251) = 4.41 $p < .037$; $\eta^2 = .017$, respectively).

Preliminary analyses revealed that there was no violation of the assumptions for normality, homoscedasticity, linearity, and independence of errors. So, a standard multiple regression was used to assess the significance of the dummy variable "affective pattern" (i.e., liking vs. disliking) predicted by the z-scores of each acoustic parameter. The variables were entered in the model by the enter method. The model ($R^2 = 0.25$, F (5, 257) = 17.188; $p < .001$) that predicted the affective patterns was

Table 3
Mean (M), standard deviation (SD) and range of the *z*-scores for acoustic parameters (mean F0, SD F0, Max F0, Min F0).

	M	SD	Min	Max
z-Score mean F0[**]				
Liking	0.40	0.74	−1.12	1.60
Disliking	−0.40	1.07	−2.06	1.86
z-Score SD F0[**]				
Liking	0.35	0.77	−1.22	2.29
Disliking	−0.35	1.08	−2.36	2.55
z-Score Max F0[**]				
Liking	0.45	0.64	−1.28	1.09
Disliking	−0.45	1.09	−2.43	1.09
z-Score Min F0[*]				
Liking	0.20	0.92	−1.66	2.37
Disliking	−0.21	1.04	−1.93	2.06

[*] $p = .001$.
[**] $p \leqslant .0001$.

composed by duration ($b = −.108$; beta = −.217), maximum F0 ($b = .246$; beta = −.490), mean F0 ($b = −.058$; beta = −.115), minimum F0 ($b = −.079$; beta = −.158), and F0 range ($b = −.016$; beta = −.032). The strongest predictor was F0 maximum, followed by duration.

As an illustration of F0 contours involved in the expression of like and dislike in EP, we plotted the mean F0 for one word. The word *limão*, lemon, was chosen for its all sonorant nature that provided a continuous F0 trace; it is a word with final stress. Liking and disliking contours differed from each other on the acoustic parameters analyzed in this study. As shown in Figs. 1 and 2, in the case of this word the strongest indicator was F0 maximum: liking had a higher F0 peak on the stressed syllable than disliking. For disliking, there was a flatter F0 pattern with a fall on the stressed syllable. Figs. 3 and 4 illustrate the typical phonological nuclear configuration of liking and disliking,
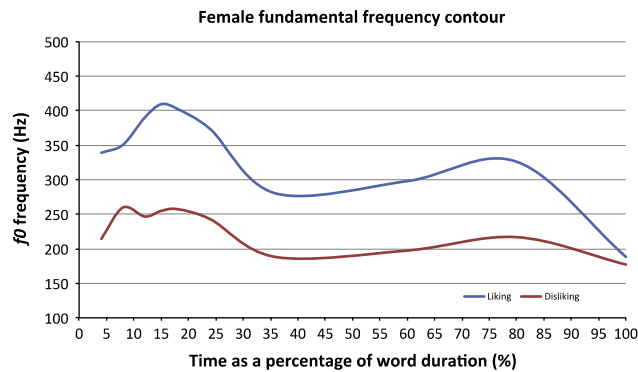
Fig. 1. Female fundamental frequency contour. Mean F0 contour of the utterances produced by female speakers ($N = 3$) for the word *limão* (lemon), for liking and disliking. In the *x*-axis, time was normalized as a percentage of the duration of the word.
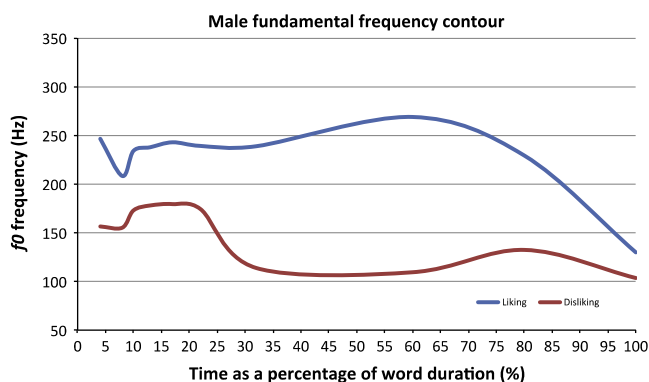

Fig. 2. Male fundamental frequency contour. Mean F0 contour of the utterances produced by male speakers ($N = 3$) for the word *limão* (lemon), for liking and disliking. In the *x*-axis, time was normalized as a percentage of the duration of the word.
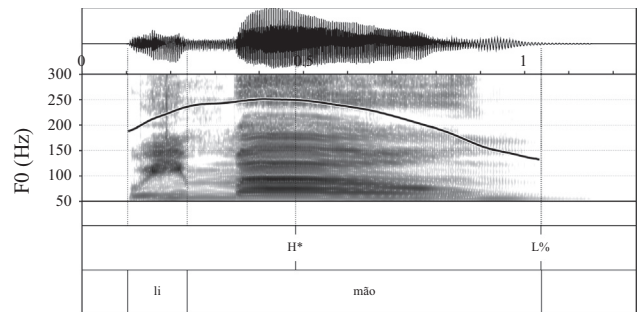

Fig. 3. F0 contour of the word *Limão* (lemon) for liking. The label tiers indicate the tonal analysis and the orthographic transcription, respectively.
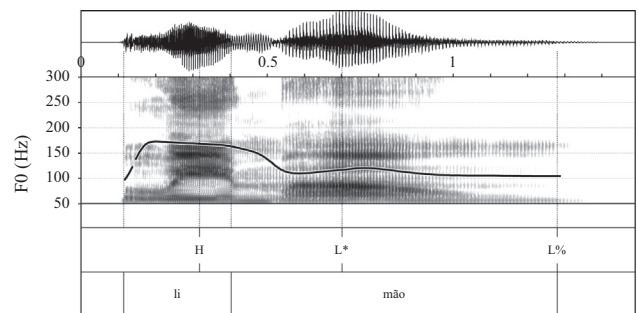

Fig. 4. F0 contour of the word *Limão* (lemon) for disliking. The label tiers indicate the tonal analysis and the orthographic transcription, respectively.

according to the autosegmental-metrical framework[1] for the analysis of intonation: the most indicative liking pattern is expressed by a high plateau on the stressed syllable followed by a fall (H* L%; 85% of the trials), whereas the most indicative disliking pattern is conveyed by a fall into the stressed syllable yielding a low plateau (H L* L%; 73% of the trials). Importantly, the stressed syllable is high in the liking pattern (H*) and low in the disliking pattern (L*). Regarding the second strongest parameter, duration, disliking was longer than liking.

4. Discussion

Given that research on affective prosody had not yet identified the vocal cues for liking and disliking in EP, the purpose of this paper was to describe and analyze the perceptual and acoustic correlates of these two patterns applied to one-word items. Our results showed that there were clear differences between the acoustic profiles of the

liking and disliking patterns, and listeners were able to identify these patterns with high accuracy. In general, the intonation feature most frequently associated with liking was a wider and higher F0 pattern, with a peak on the stressed syllable, as well as an overall rising-falling prosodic contour (H* L%). For disliking, results showed a flatter melodic pattern with a fall into the stressed syllable which displays low pitch (H L* L%). Duration was also a strong indicator, at least in one-word utterances: disliking was longer than liking. Furthermore, liking was more easily recognized than disliking.

The first research question of this study was how well liking and disliking are perceptually recognized on the basis of prosody. Our findings showed that individuals were able to identify these affective patterns with great accuracy, regardless of the speaker (73–91%). The degree of accuracy was comparable to what has been found in several studies on the identification of emotion categories using semantically neutral sentences spoken in different intonations. For instance, in Portuguese, Castro and Lima (2010) reported a mean correct of 75% for sentences and 71% for pseudosentences for the seven emotions studied. In English, the results were similar: Adolphs and colleagues (2002) reported 81% correct identifications of the five emotions studied (happy, sad, angry, afraid, and surprised). Moreover, our results suggested that the patterns of liking and disliking were not equally easy to decode: the negative pattern (disliking) was less well recognized compared to the positive one (liking). An explanation could be the use of different intonation

---

[1] The autosegmental-metrical framework uses High (H) and Low (L) tone labels to transcribe melodic targets observed on fundamental frequency curves. The label * is used for the tone associated with the stressed syllable, and the label % is used to signal a boundary tone.

features (85% of the trials presented the H[*] L% pattern for liking, whereas only 73% of the trials presented the H L[*] L% pattern for disliking) that may justify some of the misidentified stimuli in the perceptual tests. However, a similar difference between negative and positive patterns has been suggested by several authors (e.g., Adolphs et al., 2002; Banse and Scherer, 1996; Scherer et al., 1991). One possible explanation could be the fact that in natural conversation negative emotions (e.g., disgust) are probably expressed by face mimicking together with voice (Castro and Lima, 2010). Even if not all emotions have the same accuracy in identification, it is clear that humans identify emotions from speech with a very high degree of accuracy.

The second question addressed the characterization of the intonation features involved in the expression of liking and disliking in EP. Consistent with previous research in other languages (e.g., Banse and Scherer, 1996), the present experiment indicated that specific acoustic features were related to the two contrastive affective patterns studied here, and that listeners were able to reliably recognize the two patterns on the basis of vocal cues. In EP, the intonation feature most indicative of liking is a wider and higher F0 and a rising-falling contour, with high pitch on the stressed syllable. For disliking, the results showed a flatter melodic pattern with low pitch on the stressed syllable.

Finally, the last question was the cross-language comparison of the vocal profiles for affective patterns. The profile for EP is similar to English with respect to the intonation of liking – both show a rising-falling contour, but different from Spanish which shows a falling-rising contour. For disliking, Portuguese is similar to Spanish showing a flat pattern with a fall and low pitch on the stressed syllable, which is different from the fall-rise found in English (Martinez-Castilla and Peppé, 2008). Nevertheless, across languages, wider and higher pitch seems to be a general feature of liking patterns. These results suggest that the phonological contours linked to affect patterns may differ across languages (possibly, depending on the independent features of their specific intonation systems), whereas the phonetic implementation of affect patterns seems to be universal (possibly along the lines of affective interpretations of biological codes

in pitch range, as suggested in Gussenhoven, 2004). Further cross-linguistic research is needed to determine whether these suggestions are on the right track.

On the whole, this study has contributed to the characterization of the acoustic features that underlie the affective patterns of liking and disliking. Particularly, this study has revealed that both patterns are expressed in a consistently different way, but also with possible differences between languages. Future studies should address important methodological issues that remain to be explored, including studying other acoustic factors that may have an impact on the perception of liking and disliking such as intensity or voice quality, using naturalistic recordings and continuous speech, and a large variety of items (for instance, shorter and longer utterances).

Because it is socially and clinically relevant to understand how emotions are vocally expressed, the present study also has clinical implications. The main contribution – establishing the perceptual and acoustic features involved in the expression of liking and disliking in EP – offers guidelines for the understanding of affective prosody in Portuguese. As affective prosody is widely accepted as a key deficit in clinical populations such as those with schizophrenia (e.g., Pijnenborg et al., 2007), Alzheimer's disease (e.g., Taler et al., 2008), and autism (e.g., McCann and Peppé, 2003), therapists must know the normative profile in order to offer intervention, and researchers need to provide adequate assessment procedures. Therefore, the present study contributes to the characterization of affective prosody in EP with implications for its assessment and intervention in clinical populations.

Appendix A. One-word utterances included in the training and experimental sessions

Training item (Translation)	Experimental item (Translation)	Experimental item (Translation)
(1) *Noz* (Nut)	(1) *Limão* (Lemon)	(12) *Pão* (Bread)
(2) *Figo* (Fig)	(2) *Salada* (Salad)	(13) *Maçã* (Apple)
(3) *Pepino* (Cucumber)	(3) *Ervilhas* (Peas)	(14) *Iogurte* (Yogurt)
(4) *Mel* (Honey)	(4) *Ovo* (Egg)	(15) *Sumo* (Juice)
(5) *Café* (Coffee)	(5) *Laranja* (Orange)	(16) *Massa* (Pasta)
(6) *Cebola* (Onion)	(6) *Uvas* (Grapes)	(17) *Salsicha* (Sausage)
	(7) *Bolacha* (Cookie)	(18) *Banana* (Banana)
	(8) *Peixe* (Fish)	(19) *Pêra* (Pear)
	(9) *Cenoura* (Carrot)	(20) *Bolo* (Cake)
	(10) *Leite* (Milk)	(21) *Queijo* (Cheese)
	(11) *Tomate* (Tomato)	(22) *Gelado* (Ice Cream)

References

Abboud, H., Schultz, W., Zeitlin, V., 2006. SuperLab (Version 4.0) [Computer software]. Cedrus Corporation, San Pedro, CA.

Adolphs, R., Damasio, H., Tranel, D., 2002. Neural systems for recognition of emotional prosody: A 3-D lesion study. Emotion 2, 23–51. http://dx.doi.org/10.1037/1528- 3542.2.1.23.

Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. 70, 614–636. http://dx.doi.org/10.1037//0022-3514.70.3.614.

Beckman, E., Hirschberg, J., Shattuck-Hufnagel, S., 2005. The original ToBI system and the evolution of the ToBI framework. In: Jun, S.-A. (Ed.), Prosodic Typology. The Phonology of Intonation and Phrasing, Oxford University Press, Oxford, pp. 9–54. http://dx.doi.org/10.1093/acprof:oso/9780199249633.001.0001.

Boersma, P., Weenink, D., 2011. Praat: doing phonetics by computer [Computer program]. Version 5.3. <http://www.praat.org/> (retrieved 15.10.2011)

Cameirão, M.L., Vicente, S.G., 2010. Age-of-acquisition norms for a set of 1749 Portuguese words. Behav. Res. Methods 42, 474–480. http://dx.doi.org/10.3758/BRM.42.2.474.

Castro, S.L., Lima, C.F., 2010. Recognizing emotions in spoken language: a validated set of Portuguese sentences and pseudo-sentences for research on emotional prosody. Behav. Res. Methods 42, 74–81. http://dx.doi.org/10.3758/BRM.42.1.74.

Darwin, C., 1872/1965. The Expression of the Emotions in Man and Animals. University of Chicago Press, Chicago (Original work published 1872).

Duncan, L.G., Castro, S.L., Defior, S., Seymour, P.H.K., Baillie, S., Leybaert, J., Serrano, F., 2013. Phonological development in relation to native language and literacy: variations on a theme in six alphabetic orthographies. Cognition 127, 398–419. http://dx.doi.org/10.1016/j.cognition.2013.02.009.

Elfenbein, H.A., Marsh, A., Ambady, N., 2002. Emotional intelligence and the recognition of emotion from facial expressions. In: Barrett, L.F., Salovey, P. (Eds.), The Wisdom of Feelings: Psychological Processes in Emotional Intelligence. The Guilford Press, New York, pp. 37–59.

Frota, S., 2000. Prosody and Focus in European Portuguese, Phonological Phrasing and Intonation. Garland Publishing, New York.

Frota, S., 2014. The intonational phonology of European Portuguese. In: Sun-Ah, J. (Ed.), Prosodic Typology II. Oxford University Press, Oxford, pp. 6–42.

Gomes, I., Castro, S.L., 2003. Porlex, a lexical database in European Portuguese. Psychologica 32, 91–108.

Gussenhoven, C., 2004. The Phonology of Tone and Intonation. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511616983.

Hammerschmidt, K., Jürgens, U., 2007. Acoustical correlated of affective prosody. J. Voice 21, 531–540. http://dx.doi.org/10.1016/j.jvoice.2006.03.002.

Ishii, K., Reyes, J.A., Kitayama, S., 2003. Spontaneous attention to word content versus emotional tone: differences among three cultures. Psychol. Sci. 14, 39–46. http://dx.doi.org/10.1111/1467-9280.01416.

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. http://dx.doi.org/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?.

Psychol. Bull. 129 770–814. http://dx.doi.org/10.1037/0033 -2909.129.5.770.

Juslin, P.N., Madison, G., 1999. The role of timing patterns in recognition of emotional expression from musical performance. Music Percept. 17, 197–221. http://dx.doi.org/10.2307/40285891.

Kitayama, S., Ishii, K., 2002. Word and voice: spontaneous attention to emotional utterances in two languages. Cogn. Emot. 16, 29–59. http://dx.doi.org/10.1080/0269993943000121.

Kline, R.B., 2005. Principles and Practice of Structural Equations Modeling, second ed. Guilford Press, London.

Kotz, S.A., Paulmann, S., 2011. Emotion, language and the brain. Lang. Linguist. Compass 5, 108–125. http://dx.doi.org/10.1111/j.1749-818X.2010.00267.x.

Ladd, D.R., 2008, second ed. In: Intonational Phonology, vol. 119 Cambridge University Press, Cambridge.

Martinez-Castilla, P., Peppé, S., 2008. Intonation features of the expression of emotions in Spanish: preliminary study for a prosody assessment procedure. Clin. Linguist. Phonet. 22 (4), 363–370. http://dx.doi.org/10.1080/02699200801919802.

McCann, J., Peppé, S., 2003. Prosody in autism spectrum disorders: a critical review. Int. J. Lang. Commun. Disord. 38 (4), 325–350. http://dx.doi.org/10.1080/1368282031000154204.

Nygaard, L.C., Lunders, E.R., 2002. Resolution of lexical ambiguity by emotional tone of voice. Mem. Cogn. 30, 583–593. http://dx.doi.org/10.3758/BF03194959.

Paulmann, S., Pell, M.D., Kotz, S.A., 2008. How aging affects the recognition of emotional speech. Brain Lang. 104, 262–269. http://dx.doi.org/10.1016/j.bandl.2007.03.002.

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. J. Phonet. 37, 417–435. http://dx.doi.org/10.1016/j.wocn.2009.07.005.

Pijnenborg, G.H.M., Withaar, F.K., van den Bosch, R.J., Brouwer, W.H., 2007. Impaired perception of negative emotional prosody. Clin. Neuropsychol. 21 (5), 762–775. http://dx.doi.org/10.1080/13854040600788166.

Scherer, K.R., 1979. Non-linguistic indicators of emotion and psychopathology. In: Izard, C.E. (Ed.), Emotions in Personality and Psychopathology. Plenum, New York, pp. 495–529.

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

Scherer, K.R., Banse, R., Wallbott, H.G., Goldbeck, T., 1991. Vocal cues in emotion encoding and decoding. Motiv. Emot. 15, 123–148. http://dx.doi.org/10.1007/BF00995674.

Scherer, K.R., Johnstone, T., Klasmeyer, G., 2003. Vocal expression of emotion. In: Davidson, R., Scherer, K.R., Goldsmith, H. (Eds.), Handbook of the Affective Sciences. Oxford University Press, New York, pp. 433–456.

Sobin, C., Alpert, M., 1999. Emotion in speech: the acoustic attributes of fear, anger, sadness, and joy. J. Psycholinguist. Res. 28 (4), 347–365. http://dx.doi.org/10.1023/A:1023237014909.

Taler, V., Baum, S., Chertkow, H., Saumier, D., 2008. Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease. Neuropsychology 22, 188–195. http://dx.doi.org/10.1037/0894-4105.22.2.188.

Wells, B., Peppé, S., 2003. Intonation abilities of children with speech and language impairment. J. Speech, Lang. Hear. Res. 46, 5–20. http://dx.doi.org/10.1044/1092-4388(2003/001.