Processing emotional prosody in Mandarin Chinese: A cross-language comparison

Pan Liu, Marc D. Pell

School of Communication Sciences & Disorders, McGill University
Centre for Research on Brain, Language and Music (CRBLM)
Montréal, Canada
pan.liu@mail.mcgill.ca, marc.pell@mcgill.ca

Abstract

To understand how emotional prosody is processed in Mandarin Chinese and whether it differentiates from that of other languages, we conducted a perceptual-acoustic study on a set of Chinese vocal emotional stimuli and examined how they were perceived and acoustically characterized, in comparison with four other languages, English, Arabic, German, and Hindi, reported by Pell et al. [1]. Chinese pseudo-utterances spoken in seven emotions (anger, disgust, fear, sadness, happiness, pleasant surprise, and neutrality) were first identified by a group of native Mandarin speakers in a seven forced choice task, and then subjected to acoustic analyses. Results revealed that among the seven emotions, neutrality, anger, sadness, and fear tended to be recognized most accurately. Acoustic analysis demonstrated the importance of three acoustic parameters (f0 mean, f0 range, and speech rate) in characterizing vocal emotions in Mandarin. Both the perceptual and acoustic characteristics are highly similar, although not identical, to that observed by Pell et al. [1] in English, Arabic, German, and Hindi, indicating a set of universal principles in vocal emotion communication across languages.

Index Terms: Mandarin Chinese, emotional prosody, perceptual-acoustic, cross-language

1. Introduction

Humans can efficiently understand each other's emotions from speech cues without any visual information available, in cases such as telephone conversations. The vocal features of speech, including variations in pitch, loudness, speech rate, etc., which are considered as 'emotional prosody', are a universal means to communicate emotions used by speakers from different language backgrounds [2], [3].

Existing cross-language studies of emotional prosody have reported that listeners can successfully recognize the meaning of vocal emotions in a foreign language, but always with an "out-group" disadvantage (i.e., lower accuracy than that of their native language), suggesting that there is a set of shared properties encoding emotions across languages but on the other hand, discrepancies across languages exist and exposure to vocal emotions in a specific language plays a role (e.g., [4]-[6]). Acoustic studies have also demonstrated a number of consistent tendencies in the acoustic features of vocal emotions across languages (e.g., [1], [2], [7]). Thus, it will be important and interesting to further clarify the extent to which there exists a set of universal perceptual-acoustic characteristics shared across languages in vocal emotion communication, based on direct comparisons of data from different language groups.

However, the previous literature have mostly focused on Indo-European languages, while little is known about how vocal emotions are communicated in other major languages such as Mandarin Chinese. Mandarin is a Sino-Tibetan language which is spoken by more than a billion people around the world and diverges in fundamental ways from Indo-European languages (e.g., [8]-[12]). So far, little work has been done to systematically explore the perceptual-acoustic features of emotional prosody in Mandarin, with direct comparisons with other languages. By employing a well-established database of vocal emotional stimuli in Mandarin [13], this study aims to provide evidence on the perceptual-acoustic features of Chinese vocal emotions; by qualitatively comparing these data directly with four languages from different linguistic families (English, Arabic, German, and Hindi) that were examined in one of our previous studies [1], this study will shed light on to what extent there are central tendencies in how discrete emotions are communicated vocally across different languages and cultures.

To ensure the comparability of our data with that of the previous study, identical procedures of testing and analyses were adopted as those of Pell et al. [1]. In particular, emotionally-inflected "pseudo-sentences" in Mandarin (e.g., *她在一个门文上走亮*) were recognized by native Mandarin speakers in a seven-option forced-choice identification task and then subjected to acoustic analysis. Pseudo-sentences are composed of pseudo content words conjoined by real function words, rendering them meaningless but resembling the phonetic-segmental and supra-segmental properties of the target language; they have been used effectively in the literature investigating the perception of vocal tones independent of the linguistic-semantic content (e.g., [1], [5], [14]). Based on the evidence that many languages share perceptual-acoustic properties of discrete emotions, it is hypothesized that Chinese vocal emotions will show similar perceptual-acoustic patterns as those found in the four languages in comparison [1], although unique variations may also be observed.

2. Method

2.1. Participants

Twenty-four native Mandarin speakers (12 female, 12 male) with a mean age of 25.5 ± 3.3 years were recruited for the perception study. They were all students from China who learned Mandarin from birth, lived in China until at least 18 years of age, had been away from China for less than two years, and spoke English as a second language. Each participant gave written consent for the testing was compensated $10 CAD per hour for their participation.

2.2. Materials and procedure

Eight-hundred and seventy four Chinese pseudo-utterances spoken by 4 native Mandarin speakers (2 male, 2 female) in 7 emotion categories (anger, disgust, fear, sadness, happiness,

10.21437/SpeechProsody.2014-7

pleasant surprise, and neutrality), were adopted from a validated database of Chinese vocal emotional stimuli [13]. In the perceptual study, the 874 utterances were randomly combined and divided into four blocks which were presented by Superlab presentation software (Cedrus, USA) in two testing sessions, two blocks during each session. During the testing, each utterance was played once over headphones at consistent comfortable listening level, for which the participants identified which emotion was being expressed from a list of the seven categories presented on the computer screen by clicking the mouse. All participants received practice trials prior to the first block and frequent breaks during each session. All instructions were conducted in Mandarin.

2.3. Analyses

2.3.1. Perceptual-acoustic analyses

Identical perceptual-acoustic analyses were conducted as those of the study of Pell et al. [1]. For the perceptual data, the recognition rates of each of the 7 emotion categories were calculated. Acoustic analysis was performed on the 874 items and focused on three acoustic parameters that are widely employed in the literature: mean fundamental frequency (f0Mean, in Hertz), fundamental frequency range (f0Range, in Hertz), and speech rate (SpRate, in syllables per second). The values of mean f0, maximum f0, minimum f0, and utterance duration for each item were obtained in Praat [15], based on which the three parameters were calculated. Following Pell et al. [1], in order to correct for differences in a speaker's mean voice pitch and expressive range, all f0 measures (mean f0, maximum f0, and minimum f0) were normalized in relation to the individual *resting frequency* of each speaker (i.e., the average minimum f0 value of all neutral utterances produced by that speaker, see [1] for details). Therefore, for the normalized values of f0Mean and f0Range, a value of 1 for an utterance represents a 100% increase in the speaker's resting frequency, which could be compared across speakers as a proportional value. Measures of speech rate (SpRate) were calculated by dividing the number of syllables of each utterance by the duration of that utterance, in syllables per second.

2.3.2. Statistical analysis

To evaluate whether the seven emotion categories could be differentiated both perceptually and acoustically, univariate and multivariate analysis of variance (ANOVA/MANOVA) were conducted on recognition rates and acoustic measures (f0Mean, f0Range, and SpRate), respectively. In addition, a step-wise discriminant analysis was performed on the acoustic data to explore whether the seven emotion categories could be successfully classified based on the three acoustic measures.

3. Results

3.1. Perceptual data

Emotion recognition rates were calculated for each emotion category as the target emotion hit rate (% correct). A one-way ANOVA performed on recognition rates as a function of emotion category showed a significant effect of emotion category, $F(6, 154) = 26.87$, $p < .01$. Post hoc (Tukey's HSD) comparisons revealed that neutrality (86%) was recognized most accurately, followed by anger (82%), sadness (81%), fear (80%), and happiness (70%); then disgust (67%), which is significantly lower than neutrality, anger, sadness, and fear (86%; $ps < .01$). Pleasant surprise (56%) was recognized less accurately than the other six categories ($ps < .01$). Qualitative comparisons were conducted on recognition rates between Chinese and the four languages studied by Pell et al. [1]; see Figure 1 for an illustration.

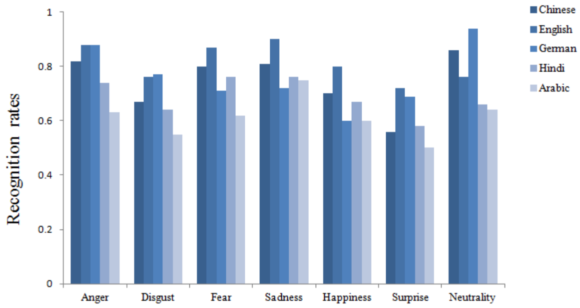

Figure 1: *Recognition rates for each emotion category of each language*.

3.2. Acoustic data

The three acoustic measures (normalized f0Mean, normalized f0Range, and SpRate) were obtained of the 874 items. To explore how the seven emotions differed in these parameters in Mandarin, a one-way MANOVA was performed on the acoustic data as a function of emotion category, with the three acoustic parameters as the dependent variables. The MANOVA indicated that the effect of emotion category on the three acoustic parameters was significant, Wilk's $\Lambda = 0.29$, $F(18, 2447) = 73.54$, $p < 0.01$. Following univariate analyses showed that the effect of emotion category was significant for f0Mean, $F(6, 867) = 106.22$, $p < 0.01$, f0Range, $F(6, 867) = 62.78$, $p < 0.01$, and SpRate, $F(6, 867) = 108.27$, $p < 0.01$. Post hoc comparisons were carried out on each acoustic parameter separately.

For f0Mean, pleasant surprise was expressed with a significantly higher f0Mean compared to all other categories ($ps < .01$), and neutrality exhibited the lowest f0Mean ($ps < .05$). Following surprise, anger and happiness yielded significantly higher f0Mean than fear ($ps < .01$), while the three emotions (surprise, anger, and happiness) showed a higher f0Mean than disgust, sadness, and neutrality ($ps < .01$). For f0Range, fewer significant differences among emotions emerged: pleasant surprise, anger, happiness, and disgust were conveyed with a significantly greater f0Range than fear, neutrality, and sadness ($ps < .01$); in addition, surprise showed a greater f0Range than disgust ($p < .01$). Finally, for SpRate, anger was expressed significantly faster than all other emotions ($ps < .01$), while disgust was produced the slowest ($ps < .01$). After anger, fear and neutrality were spoken significantly faster than happiness and sadness ($ps < .01$), while surprise and happiness also revealed a faster SpRate than sadness ($ps < .01$). Qualitative comparisons were conducted on the three acoustic parameters between Chinese and the four languages [1]. See Figure 2 an illustration.

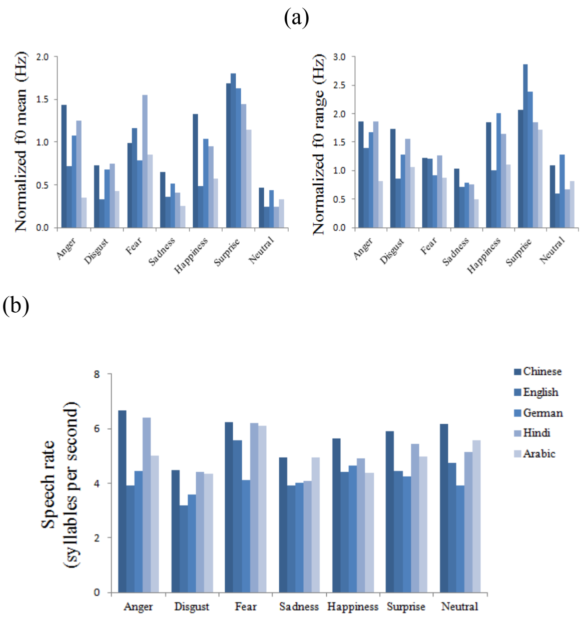

Figure 2: *(a) normalized f0 mean values for each emotion category of each language (left); normalized f0 range for each emotion category of each language (right); (b) speech rates for each emotion category of each language*.

3.3. Discriminant analysis

In order to examine how well the three acoustic parameters predicted the perceptual classification of the seven emotion categories, a discriminant analysis was performed which revealed three significant canonical functions: Function 1, $F(18, 2447) = 73.54$, $p < 0.01$; Function 2, $F(12, 1732) = 99.18$, $p < 0.01$; Function 3, $F(6, 867) = 108.27$, $p < 0.01$. Function 1 explained 55.3% of the variance and correlated significantly with f0Mean ($r = 0.80^*$) and SpRate ($r = 0.76^*$). Function 2 accounted for 38.4% of the remaining variance. Function 3 accounted for 6.3% of the remaining variance and correlated with f0Range ($r = 0.67^*$). This model successfully predicted the classification of the seven emotion categories at an overall rate of 49.9% (436/874), which is similar to those reported by Pell et al. [1] for English (58%), German (49%), Hindi (56%), and Arabic (53%) when the same analysis was performed.

4. Discussion

By conducting a perception study and acoustic analyses of emotional prosody in Mandarin which were identical to the methods of Pell et al. [1], our study allows direct comparison of the perceptual-acoustic features of emotional prosody among different languages; this allows certain inferences to be made about the extent to which vocal emotions display universal tendencies across languages.

The analysis of recognition rates showed expected variations among emotion categories, indicating that certain vocal emotions, e.g., neutrality, anger, sadness, and fear, were recognized more accurately than others in Mandarin, which is compatible with previous data in the four languages in comparison [1] and other languages such as Portuguese, Spanish, and Swedish [2], [7], [16]-[18]. A general advantage to recognize *negative* emotions from vocal speech cues, independent of language, is compatible with evolutionary views of emotion communication that vocal signals associated with threats, such as fear and anger, must be highly salient over long distances and across language systems to ensure human survival [19], [20]. While sadness is often considered a signal of the need for support from conspecifics and is therefore instrumental to maintain cohesion of social groups [21], [22], [23], [24], it is less clear from an evolutionary standpoint why sadness is often recognized most accurately in the vocal channel. Quite possibly, the recognition advantage of vocal sadness is due to their acoustic distinctiveness as speech unfolds, at least for expressions of 'depressed' sadness which lack acoustic variation and energy, as argued recently by Pell and Kotz [29].

Based on qualitative inspection of the perceptual data in Figure 1, it should be noted that fear demonstrated comparatively lower recognition rates in Arabic and German than in the other languages. One potential explanation is that the intensity of fear expressions across studies might be inconsistent; e.g., expressions of fear may be immediate and intense (i.e., 'panic' fear) or more sustained ('dread' fear) and these different forms appear to have distinct vocal cues [2] influencing recognition of this emotion. It seems that in Mandarin of this study and English and Hindi of the previous study, fear is more salient as intense portrayals of 'panic' fear with higher speech rate and f0 mean value.

Disgust and happiness were recognized relatively poorly in Mandarin compared to the other emotions, compatible with the four languages and previous findings (e.g., [2], [5], [7], [16]). For the case of happiness, it has been argued that instead of a unitary category of "happiness", there exist several positive emotions that share the facial expression of smile but possess distinct vocal cues (e.g., achievement/triumph, amusement, contentment, sensual pleasure, and relief), each of which could be reliably recognized through non-verbal vocalizations [25]. Given the fact that the unitary term 'happiness' has been used in the vocal emotion literature and the current study, it is possible that the distinct vocal cues were confounded with each other which led to difficulty in recognizing 'happy' vocal expressions.

In the case of disgust, it is more likely that in natural communication, this emotion is expressed predominantly by facial cues or by non-verbal vocalizations rather than through vocal inflections of the whole utterance (e.g., [7], [14], [26]). Note also that while disgust is considered one of the basic emotions in most studies, there is ongoing debate about whether disgust is either a sensory/interoceptive affect, or a socially constructed moral emotion; this raises the question of whether disgust would be associated with differentiated communicative properties like other basic emotions [27]. Pending further data, our findings do strongly suggest that vocal cues signifying disgust are highly distinct from other emotions in Mandarin and other languages, although the ability to recognize disgust based on prosodic cues tends to be problematic for many listeners [28], [29].

Finally, qualitative inspection showed that pleasant surprise yielded the lowest recognition rates among the seven emotion categories in Mandarin, replicating the observations of the four languages in the study of Pell et al. [1]. Interestingly, surprise was found to be recognized well in several other languages, e.g., Portuguese [7], Swedish [16], Spanish [30], and Standard Basque [31]. A significant distinction is that Pell et al. [1] and the present study elicited surprise with a positive valence ("*pleasant surprise*"), whereas all other studies elicited vocal cues conveying "surprise" without a specific valence. Compared to surprise, pleasant surprise may be more easily confounded with "happiness" and more difficult to recognize. However, as all these conclusions are based on a small group of speakers and simulated speaking contexts, it is also possible that individual biases in producing certain emotions is contributing to the variation in listeners' performance on emotion identification tasks [32].

Acoustic correlates of specific emotions

Based on qualitative inspection of the acoustic data in Figure 2, the acoustic patterns of several emotions in Mandarin demonstrated a number of consistencies with those in the four languages studied by Pell et al. [1]. Specifically, sadness was conveyed with a low f0Mean, a low f0Range (i.e., reduced variation in f0), and a slow speech rate; disgust exhibited a low f0Mean, moderate f0Range, and the slowest speech rate, whereas pleasant surprise exhibited the highest f0Mean, the largest f0Range, and a moderate speech rate. In addition, neutral speech displayed a relatively low f0Mean, narrow f0Range, and a moderate speech rate (see also [5], [7], [17], [33]-[37]). The fact that many vocal emotions are encoded acoustically in similar ways, irrespective of the language, is consistent with the idea that emotional communication is constrained to a large extent by biological factors and share a set of universal properties across languages [38].

Qualitative comparison across languages also showed greater acoustic variability in certain emotions. For example, exemplars of fear in Mandarin exhibited a moderate f0Mean whereas fear exhibited a much higher f0Mean than other emotions in English, Hindi, and Arabic in the study of Pell et al. [1]. Anger exhibited a high f0Mean whereas f0 values of anger were comparatively lower in Arabic, English, German [1]. As noted earlier, these different acoustic patterns observed of fear and anger across languages are likely to reflect two types of the target emotion. Specifically, for the case of fear, it was suggested that higher f0Mean values characterized "panic fear" while low f0Mean values characterized "sustained fear" (e.g., [2], [5]). Similarly, for anger, higher f0Mean values tend to depict hot anger (i.e., rage or intense frustration), while moderate or low f0 values tend to depict cold anger (i.e., threat; [2], [5], [39]). The possibility that speakers from different cultural-linguistic backgrounds arrive at different interpretations of the labels "fear" and "anger" that influence how they display these emotions vocally could partly explain the acoustic discrepancies observed in f0 values of these two emotions across languages.

Happiness in Mandarin was spoken with a moderate f0Mean, a moderate f0Range, and a moderate speech rate in comparison with the other emotions, whereas this emotion showed a faster speech rate than all other emotions in German [1], and much higher f0 values in Portuguese [7] and Spanish [30]. As discussed earlier, the inconsistent acoustic patterns may be due to the existence of different types of positive emotions rather than a unitary category of happiness, which were associated with different constellations of acoustic features [40]. Therefore, the "happy" expressions in the current study and the literature may actually be a mixture of these different patterns of acoustic cues, which result in discrepant results across languages. Another explanation for the acoustic variability of happiness involves different functions of negative and positive emotions. While the communication of negative emotions such as fear and anger may reflect biologically-driven responses to threat that are signaled to conspecifics in similar ways across language groups, the communication of positive emotions, which facilitate cohesion and affiliation within group, may be largely restricted to in-group members with whom the primary social connections are built and maintained [40]. Thus, as opposed to negative emotions, expressions of positive emotion are influenced by cultural rules and language variables to a larger extent (see [1], [40], [41]), and as such, display greater variability in how they are encoded and decoded across cultures and languages (see [42], [25]). In future studies, it is necessary to disentangle the unitary term "happiness" that has been used in the literature to elaborate the perceptual-acoustic features of distinct types of positive emotions. In addition, cross-cultural/linguistic studies are in need to further clarify the extent to which different types of negative and positive emotions are shaped by culture and language.

Despite focusing on only three acoustic parameters that are critical in emotional communication, a discriminant analysis of the perceptually validated items in Mandarin showed that combined changes in f0Mean and speech rate accounted for approximately 55% of the variance in the acoustic data across emotions. These results are compatible with the established view that a speaker's voice register and articulation rate are essential cues in communicating vocal emotions across languages (e.g., [1], [7], [18]). In total, the three parameters successfully predicted the classification of 49% of items into their perceived emotion categories, which is comparable to that found in other languages (English = 58%, German = 49%, Hindi = 56%, Arabic = 53%, Pell et al., 2009). Nonetheless, as approximately half of the items could not be classified by this small set of acoustic variables, future work will need to include measures of intensity/amplitude, voice quality, and other parameters to fully capture how listeners use acoustic cues to recognize emotion from speech prosody.

5. Conclusion

By conducting a perceptual-acoustic study on vocal emotion expressions in Mandarin Chinese and comparing the data directly with four other languages from a previous study (Pell et al., 2009), this study supplies new evidence of the perceptual-acoustic features of emotional prosody in Mandarin which exhibit many similar, although not identical, patterns to those found in Indo-European and Semitic languages studied by Pell et al. [1] as well as others (e.g., [2], [6], [7], [17], [18]). The new evidence from Mandarin Chinese, which is a Sino-Tibetan ideographic language with little similarity with other languages in the literature, implies that linguistic structure does not impact the core acoustic-perceptual features of vocal emotion expressions in fundamental ways and further strengthens the notion that vocal emotion communication is governed by 'universal' principles across different languages and cultures. Future research using stimuli from spontaneous speech of multiple speakers will further clarify the perceptual-acoustic features of emotional prosody in more natural day-to-day settings. Another interesting topic for future studies is to investigate how the variation of lexical tones of Mandarin influences the acoustic characteristics of emotional prosody in this language.

6. References

[1] Pell, M. D., Paulmann, S., Dara, C., Alasseri, A. and Kotz, S. A., "Factors in the recognition of vocally expressed emotions: A comparison of four languages", Journal of Phonetics, 37(4): 417-435, 2009.

[2] Banse, R. and Scherer, K. R., "Acoustic Profiles in Vocal Emotion Expression", Journal of Personality and Social Psychology, 70(3): 614-636, 1996.

[3] Bollinger, D., "Intonation: Selected readings", England: Penguin Books Ltd, 1972.

[4] Elfenbein, H. A. and Ambady, N., "On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis", Psychological Bulletin, 128(2): 203-235, 2002.

[5] Pell, M. D., Monetta, L., Paulmann, S. and Kotz, S. A., "Recognizing Emotions in a Foreign Language", Journal of Nonverbal Behavior, 33(2): 107-120, 2009.

[6] Scherer, K. R., Banse, R. and Wallbott, H. G., "Emotion inferences from vocal expression correlate across languages and cultures", Journal of Cross-Cultural Psychology, 32(1): 76-92, 2001.

[7] Castro, S. L. and Lima, C. F., "Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudosentences for research on emotional prosody", Behavior Research Methods, 42(1): 74-81, 2010.

[8] Li, K., "The information structure of Mandarin Chinese: Syntax and prosody", Dissertation Abstracts International: The Humanities and Social Sciences, 70(4): 1258, 2009.

[9] Yin, H., "The so-called Chinese VV compounds—A continuum between lexicon and syntax", Proceedings of the 2010 annual conference of the Canadian Linguistic Association, Montreal, Canada, 2010.

[10] Lin, T. H. J. and Liu, C. M. L., "'Again' and 'again': A grammatical analysis of you and zai in Mandarin Chinese", Linguistics, 47(5): 1183-1210, 2009.

[11] Gandour, J., Tong, Y., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., et al., "Neural basis of first and second language processing of sentence-level linguistic prosody", Human Brain Mapping, 28(2): 94-108, 2007.

[12] Lai, C., Sui, Y. and Yuan, J., "A Corpus Study of the Prosody of Polysyllabic Words in Mandarin Chinese", Proceedings of Speech Prosody, Chicago, USA, 2010.

[13] Liu, P. and Pell, M. D., "Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal stimuli", Behavior Research Methods, 44: 1042-1051, 2012.

[14] Paulmann, S. and Pell, M. D., "Is there an advantage for recognizing multi-modal emotional stimuli?", Motivation and Emotion, 35(2): 192-201, 2011.

[15] Boersma, P. and Weenink, D., "Praat, a system for doing phonetics by computer", Glot International, 5(9/10): 341-345, 2001.

[16] Abelin, Å., "Spanish and Swedish interpretations of Spanish and Swedish emotions – the influence of facial expressions", Proceedings of Fonetik, Stockholm, Sweden, 108–111, 2004.

[17] Abelin Å. and Allwood J., "Cross linguistic interpretation of emotional prosody", Proceedings of the ISCA Workshop on Speech and Emotion, Newcastle, Northern Ireland, 110–113, 2000.

[18] Thompson, W. F. and Balkwill, L. L., "Decoding speech prosody in five languages", Semiotica, 158: 407-424, 2006.

[19] Tooby, J. and Cosmides, L., "The past explains the present: Emotional adaptations and the structure of ancestral environments", Ethology and Sociobiology, 11(4-5): 375-424, 1990.

[20] Ohman, A., Flykt, A. and Esteves, F, "Emotion drives attention: Detecting the snake in the grass", Journal of Experimental Psychology: General, 130(3): 466-478, 2001.

[21] Barbee, A. P., Rowatt, T. L. and Cunningham, M. R., "When a friend is in need: Feelings about seeking, giving, and receiving social support", In P.A. Andersen & L.K.Guerrero (Eds.), Handbook of communication and emotion: Research, theory, applications, and contexts, 281-301, San Diego, CA: Academic Press, 1998.

[22] Miceli, M. and Castelfranchi, C., "The Plausibility of Defensive Projection: A Cognitive Analysis", Journal for the Theory of Social Behaviour, 33(3): 279-301, 2003.

[23] Sadoff, R. L., "On the nature of crying and weeping", Psychiatric Quarterly, 40(3): 490-503, 1966.

[24] Sarbin, T. R., "Emotions as narrative employments", in Entering the Circle: Hermeneutic Investigation in Psychology, 185-201, Albany, NY: State University of New York Press, 1989.

[25] Sauter, D. and Scott, S. K., "More than one kind of happiness: Can we recognize vocal expressions of different positive states?", Motivation and Emotion, 31(3): 192-199, 2007.

[26] Scherer, K. R., Banse, R., Wallbott, H. G. and Goldbeck, T., "Vocal cues in emotion encoding and decoding", Motivation and Emotion, 15(2): 123-148, 1991.

[27] Panksepp, J., "Criteria for basic emotions: Is DISGUST a primary "emotion"?" Cognition and Emotion, 21(8): 1819-1828, 2007.

[28] Jaywant, A. and Pell, M. D., "Categorical processing of negative emotions from speech prosody", Speech Communication, 54: 1-10, 2012.

[29] Pell, M. D. and Kotz, S. A., "On the time course of vocal emotion recognition", PLoS ONE, 6(11): e27256, 2011.

[30] Montero, J.M., Gutiérrez-Arriola, J.M., Colás, J., Guarasa, J.M., Enríquez, E. and Pardo, J.M., "Development of an emotional speech synthesiser in Spanish", Proceedings of Eurospeech, Budapest, Hungary, 1999.

[31] Navas, E., Hernáez, I., Castelruiz, A. and Luengo, I., "Obtaining and evaluating an emotional database for prosody modeling in standard basque", Lecture Notes in Artificial Intelligence, 393-400, 2004.

[32] Wallbott, H. G. and Scherer, K. R., "Cues and Channels in Emotion Recognition", Journal of Personality and Social Psychology, 51(4): 690-699, 1986.

[33] Iriondo, I., Guaus, R., Rodriguez, A., Lazaro, P., Montoya, N.,Blanco, J. M. et al., "Validation of an acoustical modeling of emotional expression in Spanish using speech synthesis techniques", Proceedings of the ISCA Workshop on Speech and Emotion, Newcastle, Northern Ireland, 161-166, 2000.

[34] Pell, M. D., "Influence of emotion and focus location on prosody in matched statements and questions", Journal of the Acoustical Society of America, 109(4): 1668-1680, 2001.

[35] Juslin, P. N. and Laukka, P., "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?", Psychological Bulletin, 129(5): 770-814, 2003.

[36] Liscombe, J., Venditti, J. and Hirschberg, J., "Classifying Subjective Ratings of Emotional Speech", Proceedings of Eurospeech, Geneva, Switzerland, 2003.

[37] Sobin, C. and Alpert, M., "Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy", Journal of Psycholinguistic Research, 28(4): 347-365, 1999.

[38] Scherer, K. R., "Vocal Affect Expression: A Review and a Model for Future Research", Psychological Bulletin, 99(2): 143-165, 1986.

[39] Scherer, K. R., London, H. and Wolf, J. J., "The voice of confidence: Paralinguistic cues and audience evaluation", Journal of Research in Personality, 7(1): 31-44, 1973.

[40] Sauter, D. A., Eisner, F., Ekman, P. and Scott, S. K., "Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations", Proceedings of the National Academy of Sciences of the United States of America, 107(6): 2408-2412, 2010.

[41] Rigoulot, S., Wassiliwizky, E. and Pell, M. D., "Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition", Frontiers in Psychology, 4: 1-14, 2013.

[42] Cornew, L., Carver, L. and Love, T., "There's more to emotion than meets the eye: A processing bias for neutral content in the domain of emotional prosody", Cognition and Emotion, 24(7): 1133-1152, 2010.