On the acoustic vicinity of (adult) crying and (song-like) laughing

Jürgen Trouvain
Saarland University, Saarbrücken, Germany
trouvain [at] coli.uni-saarland.de

Belin et al. [3] noticed in their database of emotional vocalisations an acoustic similarity between the samples portraying the categories "sad" (realised as crying) and "happy" (realised as laughing). This observation is in line with anecdotal evidence of many people who felt unsure whether somebody was crying or laughing when visual and other context information was missing. This is a situation which is highly irritating given the fact that crying is usually associated with negative feelings, and laughing often with positive emotions. This study has two aims: i) a comparison of selected acoustic parameters in the samples of crying and laughing in the above mentioned database [3], ii) the manipulation of cries in order to elicitate the impression of laughter.

To our knowledge there is a lack of phonetic comparisons between (adult) crying and laughing vocalisations, be it on the level of acoustics, perception or vocal production. A notable exception is the study of Erickson [5] investigating laughing, smiled and sad *speech* (not vocalisations). They too, found that "sad and happy speech were very similar in terms of acoustics and articulation" and that their "perception results showed confusion between smile and sad speech." This research gap seems to be true for adult crying in general. Leading experts on crying research such as Vingerhoets states that he is not "aware of any studies analysing the acoustical features of adult crying and the extent to which these features are similar to, or different from, child and infant crying." [11] Although there is a number of studies focusing on the acoustic characteristics of infant crying experimental studies controlling selected acoustic parameters such as F0 range are exceptions (e.g. [7]).

The data for this study were taken from the Montreal Affective Voice Database [3] with 10 actors portraying the emotional categories happiness, sadness, fear, anger, pleasure, pain, surprise, and disgust as vocalisations. Listeners selected the best examples for each speaker as prototypes. Our analysis concentrates on the 10 vocalisations (one per actor) for happy and sad, respectively (resulting in 20 vocalisations). The acoustic analysis reveals a rhythmical similarity between the expressions of both categories: in 19 of 20 vocalisations we find staccato-like quasi-"syllabic" structure [9, 10] with "happy" portrayed as song-like laughter and "sad" expressed as crying. For tempo – quantified here as "call"-rate, a "call" as equivalent to an articulated syllable [2, 9] – we see that laughter (mean 5.3 calls/sec) for all subjects is faster than crying (mean 3.4 syll/sec). In addition, laughter utterances are shorter than crying utterances for 8 out of 10 subjects (mean 1.446 sec vs. 2.229 sec). The mean fundamental frequency is higher for laughter and crying compared to "neutral". Sometimes values up to 500 Hz for male voices and up to 700 Hz for female voices are found. Although there is a tendency for crying showing a higher pitch than laughter, there are great inter-individual differences. A remarkable detail is that laughter often shows one intensity peak whereas crying often reveals two intensity peaks. Regarding mean intensity values and intensity contours no clear differences were found, however there is frequently a declination pattern for intensity and F0.

In the second part of the study the "sad" vocalisations of the database were manipulated to make them confusable with "happy" vocalisations. Informal listening tests with free answers showed that signal manipulations were successful where local tempo adapations were applied by i) shortening the duration of the consonants and vowels according to the mean duration of their

laughing counterparts and ii) keep only the more intense one of both intensity peaks. A global adapation in a linear way was not successful.

Although these preliminary results are promising for new insights of crying and laughing as acoustic phenomena it remains largely unclear in which details both vocalisation categories differ in order to maintain the huge contrast in valence as one of the important emotional dimensions. Although the presented ideas could be useful for generating laughter and crying for expressive speech synthesis [4] detailed knowledge about the effects for listeners (or users) is rather limited, e.g. the complexity of laughter as a "happy"-vocalisation [8, 9, 10] or the impression of authenticity [6]. However, laughing and crying should not be reduced to affective reflexes. There is evidence that laughter and crying are optimal carrier for memorising spoken information [1] thus bearing a potential beyond non-verbal communication.

One need for future research on adult crying and laughter is diversity of data and its elicitation. As Vingerhoets et al. [11] put it: "Crying is a rather rare behaviour, that is not easily induced in ethically acceptable ways. Work on crying would be enriched by naturalistic observations of crying behaviour." The main problem here remains the high degree of reluctance (of adults) to cry in public. Consequently it can be hard (though not impossible) to find authentic adult crying in corpora of speech or natural data from mass media. In the workshop we present examples from public TV with negative connotation (loss of the partner) as well as positive connotation (Olympic gold medal winner).

References

[1] Armony, J.L., Chochol, C., Fecteau, S. & Belin, P. 2007. Laugh (or cry) and you will be remembered: Influence of emotional expression on memory for vocalizations. *Psychological Science* 18 (12), pp. 1027-1029.

[2] Bachorowski, J.-A., Smoski, M. & Owren, M. J. 2001. The acoustic features of human laughter. *Journal of the Acoustical Society of America* 111 (3), pp. 1582-1597.

[3] Belin, P. Fillion-Bilodeau S. & Gosselin F. 2008. The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods* 40 (2), pp. 531-539.

[4] Campbell, N. 2006. Conversational speech synthesis and the need for some laughter. *IEEE Transact on Audio, Speech, Language Processing* 14 (4), pp. 1171-1178.

[5] Erickson, D., Menezes, C. & Sakakibara, K. 2009. Are you laughing, smiling or crying? *Proc. Annual Summit and Conference of the Asia-Pacific Signal and Information Processing Assoc.*, pp. 529-537.

[6] Kipper, S. & Todt, D. 2003. The role of rhythm and pitch in the evaluation of human laughter. *Journal of Nonverbal Behaviour* 27, 255-272.

[7] Protopapas, A. & Eimas, P.D. 1997. Perceptual differences in infant cries revealed by modifications of acoustic features. *Journal of the Acoustical Society of America* 102 (6), pp. 3723-3734.

[8] Sauter, D. 2010. More than happy: The need for disentangling positive emotions. *Current Directions in Psychological Science* 19, 36-40.

[9] Szameitat, D.P., Alter K., Szameitat A.J., Wildgruber D., Sterr A. & Darwin, C.J. 2009. Acoustic profiles of distinct emotional expressions in laughter. *Journal of the Acoustical Society of America* 126 (1), pp. 354-366.

[10] Trouvain, J. 2003. Segmenting phonetic units in laughter. *Proc. 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 2793-2796.

[11] Vingerhoets, A.J.J.M., Bylsma, L. & Rottenberg, J. (2009). Crying: A biopsychosocial phenomenon. In: T. Fogen (Ed.), *Tears in the Graeco-Roman world*. Berlin & New York: de Guyter, pp. 439-475.