



Available online at www.sciencedirect.com

ScienceDirect

Computer Speech and Language xxx (2013) xxx–xxx

COMPUTER
SPEECH AND
LANGUAGE

www.elsevier.com/locate/csl

Comparing the acoustic expression of emotion in the speaking and the singing voice[☆]

Klaus R. Scherer^{a,*}, Johan Sundberg^{b,2}, Lucas Tamarit^{a,1}, Gláucia L. Salomão^{b,2}

^a *Swiss Center of Affective Sciences, University of Geneva, 7, rue des Batoirs, CH-1205 Geneva, Switzerland*

^b *Department of Speech, Music and Hearing, School of Computer Science and Communication, Royal Institute of Technology, Lindstedtsvägen 24, SE-100 44 Stockholm, Sweden*

Received 26 April 2013; received in revised form 5 October 2013; accepted 14 October 2013

Abstract

We examine the similarities and differences in the expression of emotion in the singing and the speaking voice. Three internationally renowned opera singers produced “vocalises” (using a schwa vowel) and short nonsense phrases in different interpretations for 10 emotions. Acoustic analyses of emotional expression in the singing samples show significant differences between the emotions. In addition to the obvious effects of loudness and tempo, spectral balance and perturbation make significant contributions (high effect sizes) to this differentiation. A comparison of the emotion-specific patterns produced by the singers in this study with published data for professional actors portraying different emotions in speech generally show a very high degree of similarity. However, singers tend to rely more than actors on the use of voice perturbation, specifically vibrato, in particular in the case of high arousal emotions. It is suggested that this may be due to by the restrictions and constraints imposed by the musical structure.

© 2013 Elsevier Ltd. All rights reserved.

Keywords: Vocal expression; Emotional interpretation in singing; Comparison between emotion expression in speech and singing; Acoustic analyses of emotion

1. Introduction

Affect bursts (short, spontaneous, nonverbal expressions of emotion in voice, face, and body) can be considered as “living fossils” of early human affect expression that may have served as precursors of parallel evolution of speech, song, music and dance. Affect expression is likely to have played a special role because (1) innate mechanisms for spontaneous expression in face, voice, and body may have been the earliest communicative mechanisms in place, rapidly followed for control structures for learned behaviors, (2) both affect bursts and controlled vocalizations are widely shared across many species, and (3) the production mechanisms, at least for spontaneous expressions, are located in the subcortical regions of the mammalian brain (see [Scherer, 2013a,b](#)). This assumption is compatible with

[☆] This paper has been recommended for acceptance by ‘Dr. S. Narayanan’.

* Corresponding author. Tel.: +41 22 379 9211; fax: +41 22 379 9844.

E-mail addresses: klaus.scherer@unige.ch (K.R. Scherer), jsu@csc.kth.se (J. Sundberg), Lucas.Tamarit@unige.ch (L. Tamarit), gsalomao@kth.se (G.L. Salomão).

¹ Tel.: +41 22 379 9801; fax: +41 22 379 9844.

² Tel.: +46 8 790 7561; fax: +46 8 790 7854.

other suggestions concerning the origin of speech and music (such as the role of gestures or phonetic symbolism) and does not address the issue of their respective evolutionary priority (in fact, the common production substrate does not disambiguate between different evolutionary scenarios). The parallelism of human emotion expression in speech and music has been demonstrated by a comprehensive review of empirical studies on patterns of acoustic parameters in these two forms of human affect communication (Juslin and Laukka, 2003). The assumption of powerful “affect primitives” in speech and language is also supported by research on the recognition of emotion in speech (Bryant and Barrett, 2008; Laukka et al., 2013b; Pell et al., 2009; Sauter et al., 2010; Scherer et al., 2001) and music (Laukka et al., 2013a). This research has generally shown the existence of both a fairly high degree of universality of the underlying expression and recognition mechanisms and of sizeable differences between cultures, especially for self-reflective, social, and moral emotions.

The emotional power of the speaking voice, taken for granted ever since the ancient works of rhetoric (Cicero, Quintilian), has been frequently studied in empirical research (Scherer, 1986, 2003). The emotional power of the singing voice, although frequently acknowledged (Scherer, 1995; Sundberg, 1989), has only rarely been studied in an experimental fashion. Yet, in performing vocal music in Western music traditions (liturgical works, opera, and different kinds of song), professional singers have to be able to produce an extraordinary range of emotional meanings. A wide variety of means help to achieve such interpretation, such as gestures and facial expression, but the central instrument to evoke the subtle shadings of emotion is the human voice. In fact, singers are often judged in terms of their ability to produce the emotional modulation of their vocal performance that is considered appropriate to nature of the emotion to be portrayed (as implied by the text of a song or the libretto of an opera). In addition, like actors in spoken theater, they are expected to express the required emotions in a credible, authentic fashion, giving the impression of “inhabiting” the emotional feelings of the character they are performing (Scherer, 2013a).

Previous studies that have been devoted to the understanding of the emotional power of the singing voice tried to identify the acoustic cues used by listeners to recognize the emotional meaning in singing voice. A standard method for this purpose has been to correlate acoustic profiles of different emotions portrayed by professional singers with the listeners’ judgments of emotions perceived (Kotlyar and Morozov, 1976) as well as with the level of expressiveness (Sundberg et al., 1995) or strength of the perceived emotion (Jansens et al., 1997). Results have agreed in that performances characterized by higher arousal levels (as joy and anger) tend to show higher average sound pressure levels and fast tempi than performances characterized by lower arousal levels (as sadness). Also a clear association between anger and the presence of vibrato, and sadness and the absence of vibrato has been found (Jansens et al., 1997). Another method used to investigate how the emotional meaning is conveyed in singing voice has been to compare recordings of a same song performed by various singers and analyze listeners’ judgments of emotions perceived for each particular performer. Siegwart and Scherer (1995) and Howes et al. (2004) found that listeners’ preferences and emotion judgments were indeed associated with specific acoustics characteristics. Correlations between different acoustic parameters and listeners’ perception of emotions in singing voice (as well as in music in general) can also be studied by investigating the listeners’ emotion judgments of sounds that had each of different parameters systematically and independently manipulated (Scherer and Oshinsky, 1977; Kotlyar and Morozov, 1976). Procedures of synthesis and resynthesizes have also been used to systematically manipulate acoustics parameters, in order to investigate the effects and relevance of each of the parameters for listeners emotion judgment (e.g., Goto et al., 2012; Fonseca, 2011; Kenmochi and Ohshita, 2007; Risset, 1991; Sundberg, 1978). A comparison between acoustic patterns that characterizes both expressive speech and expressive singing suggests a striking parallel between the expression of emotions in the speaking and the singing voice between. For instance, both in speech and singing anger is associated with high F0 variability (assuming that F0 variability in speech is translated into vibrato extent in singing) and with high vocal intensity, while sadness is associated with slow speech rate (and low tempo) as well as with low vocal intensity.

In consequence, we expect that emotion expression is similar in speaking and singing voice – because of the evolutionary origin of the expression mechanisms and the need for authenticity (Maynard Smith and Harper, 2003; Mortillaro et al., 2013). Obviously, there may well be important differences across languages and cultures due in large part to language characteristics such as phonemic structure or intonation rules. Yet, given the stability of findings across music and speech (Juslin and Laukka, 2003), one can expect similarities across studies in different languages and cultures between the expression of emotion in speech and singing.

Unfortunately, this issue not well researched. Compared to the study of facial expression of emotion, research on vocal expression is relatively rare, particularly with respect to the comparison between languages and cultures (see a recent review by Scherer et al., 2011). To the best of our knowledge, there have been no systematic, empirical attempts

Download English Version:

<https://daneshyari.com/en/article/10368593>

Download Persian Version:

<https://daneshyari.com/article/10368593>

[Daneshyari.com](https://daneshyari.com)