Vocal emotions on the brain: The role of acoustic parameters and musicality

Summary of the dissertation by Christine Nussbaum

Background and research questions:

The human voice is an important transmitter of social signals, even beyond the spoken word. Our emotions in particular become audible through the sound of our voice and the ability to perceive emotions from human voices accurately and efficiently is of great importance in everyday interaction. Therefore, understanding how **we perceive emotions in the voice and what individual differences there are in this ability** is of great scientific interest.

Previous research has shown that human listeners are remarkably good at inferring emotions from the human voice. This recognition performance is based on highly efficient and automatic processing of **acoustic voice cues**, which give away our emotional states. When we are angry, for example, our voice usually becomes higher in pitch, significantly harsher in tone and louder. In my work, I used modern voice morphing technology to precisely control the emotional expression via these individual voice cues, allowing me to quantify their causal influence. Thus, the first question of my work was: (1) What is the contribution of different acoustic cues to vocal emotion perception?

In a second step, I looked at individual differences, with a specific focus on auditory expertise and **musicality**. The literature shows that musicians are slightly better at recognizing emotions in the voice than non-musicians. However, the mechanisms underlying this performance difference are insufficiently understood. I was particularly interested in the role auditory sensitivity and wanted to understand to which extent musicians may use acoustic cues differently or more effectively to infer vocal emotions. My second research question was therefore: (2) How does musicality affect the processing of emotional voice cues?

Finally, this thesis presents a critical reflection on the **voice morphing technology**, which was used for the first time for the acoustic manipulation of emotions and in combination with electroencephalogram (EEG) measurement in the context of this doctoral thesis. Although this technology has enormous potential for the systematic investigation of emotional voice cues, I noticed that the manipulated voices sometimes sound distorted and potentially unnatural, which can limit their ecological validity. I addressed this problem in my third question: (3) Is parameter-specific voice morphing a suitable tool to study the processing of vocal emotions?

Research methods

The methodological backbone of my dissertation is the **parameter-specific voice morphing**. I used this technology to create short vocal utterances expressing four emotions (happiness, pleasure, fear, and sadness) only via the fundamental frequency contour (F0, i.e. voice pitch), via the timbre (i.e. voice quality), or via both. I used these stimuli for behavioral experiments as well as two paradigms using electroencephalography (EEG).

My three research questions are addressed in five chapters of my dissertation which are all published in peer-reviewed journals by now:

1. a **systematic review** on the link between musicality and vocal emotion perception.
2. an empirical **online study** to validate the stimulus material which was created via parameter-specific voice morphing.
3. an empirical **EEG-study** incorporating behavioral and electrophysiological data, exploring the role of F0 and timbre for vocal emotion perception and their related ERP-modulations.
4. an empirical **online study** comparing a group of professional musicians to a group of non-musicians with regard to their vocal emotion perception performance. Of specific interest

was how musicians make use of different acoustic cues to perceive emotions, and in what way such usage might differ from that in non-musicians.

5. an empirical **EEG-study** exploring how the behavioral differences between professional musicians and non-musicians are reflected at the brain level

Results

(1) What is the contribution of different acoustic cues to vocal emotion perception?

I found that **both the F0 (i.e. pitch contour) and timbre (i.e. voice quality) convey emotional information**, which listeners can use flexibly to infer the emotion being expressed. However, their specific contribution depends on the emotion: for emotions of greater intensity, such as happiness or fear, F0 plays a much more important role than timbre. For less intense emotions, such as sadness or pleasure, the influence of pitch and timbre is more balanced. This pattern is also reflected at the neural level in the electrophysiological response.

(2) How does musicality affect the processing of emotional voice cues?

First, I was able to replicate that musicians outperform non-musicians at recognizing emotions from voice cues. Crucially, however, the data point towards a particular **importance of pitch cues for musicians**: musicians performed better than non-musicians when emotions were expressed solely through pitch, but not when they were expressed through timbre. In addition, I found that the link between auditory sensitivity and vocal emotion perception performance persisted in the absence of formal musical training, suggesting a strong role of a predisposition in individuals to exploit melodic patterns in both music and voices.

(3) Is parameter-specific voice morphing a suitable tool to study the processing of vocal emotions?

While the data showed that voice morphing can disrupt the acoustic quality of voices, perception of emotions proved to be remarkably robust against these distortions. Overall, this work therefore presents convincing evidence that **voice morphing is a valid tool** for researching vocal emotions when used with a critical awareness of its limitations.

Relevance in the field

In summary, this dissertation expands the conceptual and empirical knowledge of vocal emotion processing, regarding the underlying acoustic parameters, electrophysiological correlates, and individual differences with a specific focus on musicality. This way, it contributes to the understanding of several “common everyday” and yet extraordinary qualities in humans: **the use of our voices, the ability to express and perceive emotions, and the capacity to make music.**