Posed and spontaneous nonverbal vocalizations of positive emotions: Acoustic analysis and perceptual judgments

Roza G. Kamiloğlu, Disa A. Sauter
University of Amsterdam
r.g.kamiloglu@uva.nl

Abstract

When experiencing different positive emotional states, like amusement or relief, we may produce nonverbal vocalizations such as laughs and sighs. In the current study, we describe the acoustic structure of posed and spontaneous nonverbal vocalizations of 14 different positive emotions, and test whether listeners ($N = 201$) map the vocalizations to emotions. The results show that vocalizations of 13 different positive emotions were recognized at better-than-chance levels, but not vocalizations of being moved. Emotions varied in whether vocalizations were better recognized from spontaneous or posed expressions.

The human voice is an important carrier of emotional meaning. In addition to semantic information and speech prosody, emotions in the voice can be expressed using nonverbal vocalizations like screams, laughs, and sighs. To date, most research investigating vocal expressions of emotions has included a very limited number of positive emotions, or has used a single positive emotion, happiness, as an umbrella construct (Sauter & Scott, 2007). However, we can produce nonverbal vocalizations of a range of different positive emotional states like amusement, relief, or pleasure (e.g., Anikin & Persson, 2017; Laukka et al., 2016).

In a recent comprehensive review, we mapped the acoustic profiles of vocal expressions of all positive emotions that have been studied (Kamiloğlu et al., 2020). The results demonstrated differentiations between positive emotions, as well as between clusters of positive emotions, suggesting that different positive emotional states may correspond to distinct vocal signatures.

However, most previous research has included only a limited number of positive emotions and has tended to employ posed emotional expressions. In the current study, we examine the acoustic features characterizing nonverbal vocalizations of 14 different positive emotions, comparing posed and spontaneous expressions. In addition, we test whether these vocalizations are mapped to the corresponding specific positive emotional states by naïve listeners.

1 Posed vocal expressions

Arguably the most commonly used method for collecting emotional expressions is to ask actors or lay people to produce nonverbal vocalisations on demand (e.g., Banse & Scherer, 1996), but such posed expressions are considered lower in terms of ecological validity compared to spontaneous expressions (Williams & Stevens, 1981). By definition, posed expressions are produced voluntarily and there is little doubt that they are likely to be influenced by prototypical representations (Scherer & Bänziger, 2010). Despite these shortcomings, posed expressions provide experimental control that allows precise measurements of acoustic features (Frank et al., 2005) and certainty of the intended emotion being expressed.

2 Spontaneous vocal expressions

Realistic, spontaneous vocalizations can be taken from real-world settings like radio interviews or recorded conversations (Jürgens et al., 2011). Spontaneous vocal expressions have the advantage of reflecting genuinely felt emotions. However, recording spontaneous vocalizations of emotions is challenging and often results in low quality recordings. Nevertheless, spontaneous emotional

expressions have the advantage of being higher in terms of ecological validity compared to posed expressions.

3 The present study

In the present study, we examine a wide range of positive emotions expressed via nonverbal vocalizations. We describe the acoustic structure of vocalizations of 14 different positive emotions, and test whether naïve listeners can recognize these emotions from nonverbal vocalizations.

We included vocalizations that were posed by lay people, as well as vocalizations that were spontaneously expressed in real-world contexts. In an attempt to be inclusive of a wide range of positive emotions, a total of 14 positive emotions were included: admiration, amusement, awe, being moved, determination, elation, excitement, interest, lust, positive surprise, relief, sensory pleasure, tenderness, and triumph. Definitions of these emotions can be found in Appendix A.

4 Materials and Methods

4.1 Ethics statement

The collection of posed vocalisations and the recognition experiment were approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences, University of Amsterdam. All participants provided informed consent before participation.

4.2 Participants in perception study

A total of 201 participants (165 females, $M_{age} =$ 20.30, $Sd_{age} = 2.82$, range = 18-36 years old) were recruited via University of Amsterdam, Department of Psychology's research pool. All participants reported having no hearing impairments. The experiment session took approximately 15 minutes, and participation was compensated with monetary reward or course credit.

4.3 Materials and procedure

Posed vocalizations: The posed vocalizations were taken from a larger corpus recorded in the University of Amsterdam's psychology laboratory. Twenty participants (10 female, 10 males; $M_{age} =$ 22.42 $SD = 2.64$) whose native language was Dutch and who had never been diagnosed or treated for any voice, speech, or language disorder, were invited to the lab to record vocalizations.

Upon arriving at the lab, participants were shown each emotion term in turn on a lab computer together with its definition (see Appendix A). Participants were then asked to describe the emotion in their own words to the experimenter to ensure that they had understood the definition correctly. They then read a situational example presented on the computer screen and were asked to imagine the situation as vividly as possible. They then produced a nonverbal vocal expression of the corresponding emotion.

Participants produced nonverbal vocalisations for each target emotion. If they produced multiple vocalisations of a given emotion, they were asked to choose the expression they thought best expressed the target emotion. All stimuli were recorded using a Sennheiser MKE 600 microphone and a Tascam DR-100 MK3 recorder.

For the purpose of this study, 28 vocalisations reflecting 14 positive emotions (one male and one female vocalisation for each emotion) were randomly selected from this larger corpus, in order to match the number of spontaneous vocalisations (see below).

Spontaneous vocalizations: Spontaneous stimuli were collected from online sources (e.g., youtube.com) by two naive research assistants. The assistants searched for nonverbal vocalizations of emotions based on definitions (see Appendix A) and were free to use any search terms relating to these. Inclusion was based exclusively on 1) the eliciting context matching the definition of the positive emotion category, and 2) the clear audible presence of a nonverbal vocalization.

Additionally, the research assistants rated the videos on suddenness, clarity, and certainty (see Anikin & Persson, 2017) as well as perceived intensity, reflecting how intensely the target emotion was thought to be experienced. The videos that scored highest on these four ratings were used as stimuli, resulting in 28 nonverbal vocalizations, two for each emotion category, with the stimuli in each category balanced for gender.

4.4 Recognition study: Experimental Procedure

The study was run online using the Qualtrics survey tool. After being informed of the general procedure, listeners read the definitions of the 14 positive emotions and made a forced-choice

emotion categorization judgment ("Which emotion is expressed in this vocalization?") for each of the 56 vocalizations, selecting from 14 emotion categories. In addition to the vocalization stimuli, we added two trials with sounds as an attention check. The experiment took around 15 minutes.

The response options were presented in alphabetical order. The presentation order of the stimuli was randomized for each participant. Participants were able to replay each sound as many times as needed to make a judgment for each vocalization.

5   Statistical Analysis

To test whether listeners performed better than the chance level in recognition of emotion categories, the proportion of correct responses was calculated for each participant and emotion category. Unbiased hit rates (Wagner, 1993) were calculated to control for individual biases in the use of certain emotion categories. These were arcsine transformed and compared with the chance level (1/14) using t-tests (Bonferroni corrected for multiple testing). To test whether there were differences between recognition of posed and spontaneous expressions of positive emotions, we also compared Hu scores per emotion across posed and spontaneous expressions using paired sampled t-test with Bonferroni correction.

6   Results

6.1   Recognition of positive emotions from the voice

Confusion matrices showing listeners' categorization data per emotion category are shown in Figure 1.

The results comparing recognition of each emotion to chance level are provided in Table 1. Participants were able to accurately categorize all positive emotions except vocalizations of being moved. When compared with posed vocalizations, spontaneous vocalizations of admiration ($t = 3.81$), excitement ($t = 4.774$), and interest ($t = 8.594$) were recognized better, while vocalizations of determination ($t = -4.075$), relief ($t = -10.295$), sensory pleasure ($t = -4.560$), and tenderness ($t = -6.428$) were recognized significantly worse.
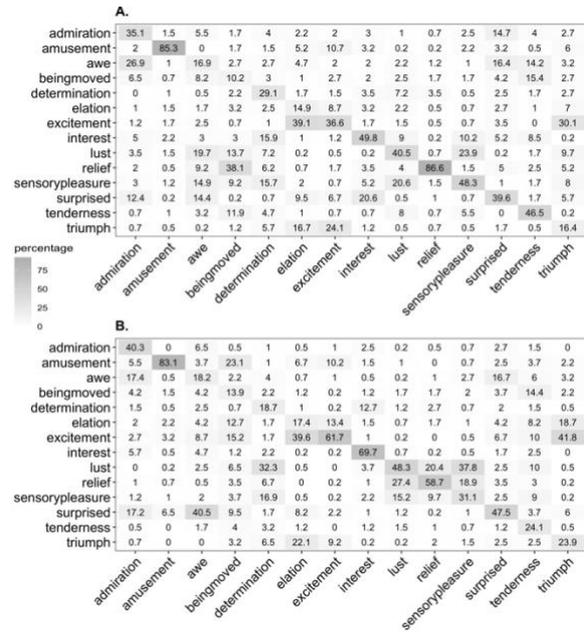

Figure 1: Confusion matrices showing percentages of responses per emotion category for Posed (A) and Spontaneous (B) vocalizations.

	t (posed)	t (spontaneous)
Admiration	**10.863**	**12.402**
Amusement	**28.261**	**27.546**
Awe	**4.984**	**5.509**
Being moved	2.447	2.247
Determination	**12.291**	**6.533**
Elation	**5.201**	**5.024**
Excitement	**11.1**	**19.532**
Interest	**13.789**	**21.645**
Lust	**14.004**	**16.486**
Relief	**29.613**	**21.264**
Sensory pleasure	**13.745**	**10.04**
Surprised	**10.15**	**13.079**
Tenderness	**12.254**	**8.934**
Triumph	**4.06**	**6.587**

Table 1: Results of t-tests comparing Hu scores against chance (1/14), using Bonferroni correction. Bold marks better-than-chance level recognition.

7   Acoustic Analysis

We performed an acoustic analysis to examine the acoustic features characterizing the emotional vocalizations. First, we extracted 88 acoustic features by utilizing the extended version of the Geneva Minimalistic Acoustic Parameter Set (GeMAPS, see Eyben et al., 2016). Second, we

performed a Principal Component Analysis (PCA) with varimax normalised rotation to reduce the number of acoustic features based on the highest factor loadings.

Based on the results of PCA on the 88 parameters and examination of the scree plot, the first four components, which together explained 44% of the variance, were retained. The acoustic features with the highest factor loadings on the first four acoustic dimensions can be found in Table 2. An online map showing the distribution of the 14 positive emotions for posed vocalizations is available from https://emotionwaves.github.io/posed14/ and spontaneous expressions from https://emotionwaves.github.io/spontaneous14/. Example vocalizations of each positive emotion can be listened from the same links.

	Acoustic features	Factor loadings
Component 1	F0 Mean	.82
	F0 Percent Range	.86
	Mfcc2	.62
Component 2	Loudness Percent	.82
	Spectral Flux Mean	.81
Component 3	F1 Amplitude	.62
	Hammarberg	.62
	Harmonics-to-Noise Ratio	.67
	Mfcc1	.62
Component 4	Spectral Flux SD	.57

Table 2: Acoustic features with highest loadings on the first four components.

8   Discussion

We present an acoustic map of posed and spontaneous nonverbal vocalizations of 14 positive emotions. Listeners were accurate at recognizing 13 positive emotions from both posed and spontaneous nonverbal vocalizations in a forced-choice task. Being moved was the only emotion category that was not recognized from the vocalizations. Overall, these results suggest considerable granularity in the perception of nonverbal vocalizations of positive emotions.

There was considerable variability in recognition levels of the different types of vocalizations. Some emotions, including relief and amusement, were recognized particularly well. A possible explanation is that these emotions have strong perceptual anchors, prototypical vocalizations, reflected in distinct acoustic patterns. Some emotions were frequently confused with each other, such as triumph and excitement (see Figure 1). The confusion between these emotions might be due to similarities of their acoustic profiles.

Recognition levels per positive emotion were different for posed and spontaneous vocal expressions. There is unlikely to be single explanation for these differences, but differences in recording quality, emotion intensity, or cultural effects may have contributed. Moreover, the spontaneous vocalizations were selected by two judges in our study and it cannot be comprehensively ruled out that their selection could have been influenced by some degree of implicit bias. For instance, it is possible that the judges selected prototypical vocalizations of emotions, even though they were specifically instructed to choose vocalizations based only on the relevant context and definition. Furthermore, the intense expressions in the present study may differ from many emotional expressions that occur in daily life. Nevertheless, these results provide evidence that naïve listeners can recognize 13 different positive emotions from posed and spontaneous vocalizations.

Acknowledgements

References

Andrei Anikin and Tomas Persson. 2017. Nonlinguistic vocalizations from online amateur videos for emotion research: A validated corpus. *Behavior Research Methods,* 49, 758–771. https://doi.org/10.3758/s13428-016-0736-y

Rainer Banse and Klaus R. Scherer, 1996. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology,* 70, 614. https://doi.org/10.1037/0022-3514.70.3.614

Florian Eyben, Klaus R. Scherer, Björn W. Schuller,… Khiet P. Truong. 2016. The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for voice research and affective computing. *IEEE Transactions on Affective Computing,* 7, 190–202. https://doi.org/10.1109/TAFFC.2015.2457417

Mark G. Frank, Patrik N. Juslin, and Jinni A. Harrigan. 2005. Technical issues in recording nonverbal behavior. In J. A. Harrigan, R. Rosenthal, & K. R.

Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 449–470). New York: Oxford University Press.

Rebecca Jürgens, Kurt Hammerschmidt, and Julie Fischer. 2011. Authentic and play-acted vocal emotion expressions reveal acoustic differences. *Frontiers in Psychology,* 2, 1–11. https://doi.org/10.3389/fpsyg.2011.00180

Roza G. Kamiloğlu, Agneta H. Fischer, and Disa A. Sauter. 2020. Good vibrations: A review of vocal expressions of positive emotions. *Psychonomic Bulletin & Review,* 1-29. https://doi.org/10.3758/s13423-019-01701-x

Petri Laukka, Hillary A. Elfenbein, Nutankuman S. Thingujam, Thomas Rockstuhl, Fredrick K. Iraki, Wanda Chui, Jean Althoff. 2016. The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features. *Journal of Personality and Social Psychology, 111*, 686. https://doi.org/10.1037/pspi0000066

Disa A. Sauter. and Sophie Scott. 2007More than one kind of happiness: Can we recognize vocal expressions of different positive states?. *Motivation and Emotion,* 31, 192-199. https://doi.org/10.1007/s11031-007-9065-x

Klaus R. Scherer and Tanja Bänziger, T. (2010). On the use of actor portrayals in research on emotional expression. In K. R. Scherer, T. Bänziger, & E. Roesch (Eds.), *A blueprint for an affectively competent agent: Cross- fertilization between emotion psychology, affective neuroscience, and affective computing* (pp. 166–176). Oxford, UK: Oxford University Press.

Hugh L. Wagner. 1993. On measuring performance in category judgment studies of nonverbal behaviour. *Journal of Nonverbal Behaviour,* 17, 3-28. https://doi.org/10.1007/BF00987006

Carl E. Williams and Kenneth N. Stevens, K. N. (1981). Vocal correlates of emotional states. In J. K. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 221–240). New York, NY: Grune and Stratton.

Supplementary Materials

Appendix A. Definition of positive emotions

1. Admiration
The feeling when you look up to someone who has excellent abilities or has accomplished impressive things. You have the urge to also achieve such things and to be more like this person.

2. Amusement
The feeling when you encounter something silly, ironic, witty, or absurd, which makes you feel entertained. You have the urge to be playful and share the joke with others.

3. Awe
The feeling when you encounter something that is greater or more powerful than yourself. You feel a sense of overwhelming positivity and need a moment to adjust.

4. Being Moved
The feeling when you encounter something very beautiful, meaningful, or bittersweet. Tears well up in your eyes and you feel overcome with warm feelings.

5. Determination
The feeling when you are working on a demanding problem or task that you believe you can handle. You feel stimulated to accomplish it.

6. Elation
The feeling when something very good just happened to you. You have the urge to celebrate and share it with others.

7. Excitement
The feeling when you expect that something good or nice will happen to you. You cannot stop thinking about this.

8. Interest
The feeling when you encounter something new and relevant that you do not immediately understand. You have the urge to find out more about it.

9. Lust/Sexual Desire
The feeling when you think about or interact with someone that you find sexually attractive. You have the urge to be physically close and have sexual relations with this person.

10. Positive Surprise
The feeling when you realize something good or nice just happened, which you did not expect. You need a moment to take in the good news.

11. Relief
The feeling when an unpleasant experience is finally over, or when you find out that something

you had dreaded has not happened or will not happen. You can finally take your mind off it.

12. Sensory Pleasure
The feeling when something happens that pleases your senses.

13. Tenderness
The feeling when you encounter someone or something that seems cute, vulnerable, or childlike in appearance or behaviour. You have the urge to nurture and care for this person or thing.

14. Triumph
The feeling of release and a great joy, after a successful ending of a struggle or contest.