The Voice Conveys Specific Emotions: Evidence From Vocal Burst Displays

Emiliana R. Simon-Thomas, Dacher J. Keltner, Disa Sauter, Lara Sinicropi-Yao, and Anna Abramson
University of California, Berkeley

Studies of emotion signaling inform claims about the taxonomic structure, evolutionary origins, and physiological correlates of emotions. Emotion vocalization research has tended to focus on a limited set of emotions: anger, disgust, fear, sadness, surprise, happiness, and for the voice, also tenderness. Here, we examine how well brief vocal bursts can communicate 22 different emotions: 9 negative (Study 1) and 13 positive (Study 2), and whether prototypical vocal bursts convey emotions more reliably than heterogeneous vocal bursts (Study 3). Results show that vocal bursts communicate emotions like anger, fear, and sadness, as well as seldom-studied states like awe, compassion, interest, and embarrassment. Ancillary analyses reveal family-wise patterns of vocal burst expression. Errors in classification were more common within emotion families (e.g., ‘self-conscious,’ ‘pro-social’) than between emotion families. The three studies reported highlight the voice as a rich modality for emotion display that can inform fundamental constructs about emotion.

Keywords: emotion, expression, vocal burst, vocalization

Empirical studies of emotion signaling in the face, voice, and touch have figured prominently in the central debates in affective science. Controversies surrounding the degree to which emotions are universal (Ekman, 1992; Elfenbein & Ambady, 2002), the reliability with which emotions are signaled in distinct communicative behaviors, (Fridlund, 1994; Russell, Bachorowski, & Fernandez-Dols, 2003) and whether less-studied states, such as embarrassment or pride, have the temporal and signal features of well-studied emotions (Keltner & Buswell, 1997; Tracy & Robins, 2004) attest to the central role that data on emotion signaling has played in an understanding of emotion.

It is widely assumed that emotions have distinct signals, which coordinate adaptive interactions between conspecifics (Keltner & Kring, 1998). Closer analysis of these signals, in turn, informs an evolutionary understanding of the origins of each signal and its accompanying state. Comparisons between nonhuman primate laughter and smiling-like behavior, for example, strongly suggest that laughter and smiling—and their accompanying states amusement and joy—have distinct origins and functions (Preuschoft & van Hooft, 1995).

The human voice is a rich and nuanced source of emotional signaling, as evident in studies of the varieties of laughter (Bachorowski, Smoski, & Owren, 2001), teasing (Keltner, Capps, Kring, Young, & Heerey, 2001), and motherese (Fernald, 1992). Nearly half of Darwin’s descriptions of the nonverbal correlates of over 40 emotion-related states include references to specific paralinguistic vocalizations—“snorts” of contempt, “little coughs” of embarrassment, “air sucks” of high spirits, and “deep sighs” of grief (Darwin, 1872). Here, we present three studies that examine how well the human voice communicates 22 different emotions. The 22 emotions studied here do not represent a definitive nor exhaustive set of emotions with specific vocalizations. Rather, they represent a collection of states that have been examined in recent studies of face or voice, suggesting themselves as states worthy of empirical inquiry here (Matsumoto, Keltner, Shiota, O’Sullivan, & Frank, 2008; Sauter & Scott, 2007).

Empirical studies of emotion vocalization generally follow one of two strategies (Scherer, Johnstone, & Klasmeyer, 2003). In a first, ‘posers’ recite neutral or nonsense phrases in a tone of voice, or affective prosody, that expresses a specified emotion (Banse & Scherer, 1996). Then, these vocal stimuli are presented to ‘judges,’ who attempt, most typically in forced choice format, to identify the emotion conveyed. A review of 39 such studies concluded that there are distinct prosodic acoustic markers for five different emotions: fear, anger, sadness, happiness, and tenderness (Juslin & Laukka, 2003).

A second strategy is to examine how the voice signals emotion not during speech, but during vocal bursts, which are brief, nonword utterances that arise between speech incidents. Vocal bursts include involuntary, emotion-laden, sounds like shrieks, groans, or grunts, as well as conventionalized, emblematic expressions like the enthusiastic ‘yeeeee-haaaw’ or the breathy, awe-inspired ‘whaaaaaaa,’ which can potentially translate to words (Scherer, 1994). In relevant research, one vocal burst study reported 81.1% accuracy for recognizing displays of 10 different emotions: admi-
ration, anger, boredom, contempt, disgust, elation, relief, threat, startle, and worry (Schroder, 2003). Another reported 70.1% accuracy for vocal bursts for five different positive emotions: achievement, amusement, contentment, pleasure, and relief, across two cultural groups (Sauter & Scott, 2007).

In light of the emergent interest in positive emotion (Fredrickson, 2001) and self-conscious emotion (Tangney & Fischer, 1995), in the present investigation we asked whether 22 different emotions could be signaled reliably with vocal bursts. More specifically, we examined whether vocal bursts communicate: (1) well-studied, negative emotions, anger, disgust, fear, sadness, and surprise, (2) a less-studied negative emotion, contempt (Matsuo & Ekman, 2004), (3) the self-conscious emotions of embarrassment, shame, and guilt, which have been differentiated in lexical studies and in terms of facial and postural data (Keltner & Buswell, 1997; Tangney & Fischer, 1995), and (4) 13 positive emotional states, defined by approach and positive valence. Given the heterogeneity of vocal bursts, from the spontaneous to the conventionalized and emblematic, we recognize that the vocal burst stimuli we gathered only partially capture the breadth of vocal bursts that are likely to covary with each of the emotions studied.

Following previous conceptual analysis and empirical data, we focused our analysis of positive emotion vocal bursts on four potential families of positive emotion. One approach would be to cluster the 13 positive states into two of four families according to their placement upon the dimensions of valence and arousal, that is, low versus high arousal positive states (J. A. Russell, 1980). We chose a different route, clustering the positive emotions into emotion families according to similarities in their theorized appraisal themes and action tendencies. Contentment, desire, and sensory pleasure, for example, all fall into the family of ‘savoring’ emotions (Ekman, 1992).

Vocal bursts for four of the positive emotions considered here fall into the ‘epistemological’ family; these are states that accompany shifts in an understanding of, or knowledge about the world (Shiota, Keltner, & John, 2006). Here, amusement (Ruch, 1993), awe (Keltner & Haidt, 2003), interest (Marshall-Reeve, 2004) and relief were considered within the ‘epistemological’ family of emotions. The second family of vocal bursts considered is ‘pro-social,’ which includes expressions of emotions that orient people toward the welfare of others like compassion, gratitude, and love (Haidt, 2003; McCullough, Kilpatrick, Emmons, & Larson, 2001). The third family of interest is ‘savoring;’ this includes states that involve thinking about, or experiencing, enjoyable stimuli, that is, contentment, desire, and sensory pleasure (Berridge & Robinson, 2003; Fredrickson, 1998). Finally, we examined an ‘agency-approach’ emotion family, associated with approach motivation related to past, present or anticipated actions that have led or will lead to improved rewards, both social (e.g., elevated status) and nonsocial (e.g., nutrients). Vocal bursts from the ‘agency-approach’ emotion family evaluated here include enthusiasm, pride, and triumph (Tracy & Robins, 2004). Given the preliminary nature of this line of inquiry, these families are most certainly arbitrary; clearly the emotions might cluster in other coherent ways. For example, love and desire could be clustered as a ‘romantic’ emotion family, awe and gratitude could be clustered as a ‘spiritual’ family, and so on.

Our investigation focused on two levels of analysis. First, we examined the extent to which 22 different emotions could be communicated with vocal bursts by naïve participants. We chose ordinary participants rather than trained posers or actors on the assumption that they would generate more ecologically valid vocal bursts, and thus offer a more representative assessment of the extent to which vocal bursts can communicate different emotions (Russell, 1994). Second, in light of conceptual analyses briefly reviewed above, we asked whether the signal value of different vocal bursts—that is, the extent to which they are reliably decoded by judges—cluster into the five emotion families of interest: ‘self-conscious,’ ‘epistemological,’ ‘pro-social,’ ‘savoring,’ and ‘agency-approach.’

**Study 1: Judgment of Vocal Bursts for Negative and Self-Conscious Emotions**

**Method**

**Vocal burst stimuli.** Emotion vocal burst stimuli were gathered from 10 untrained posers (six male) during individual, 45- to 60-min appointments. Each poser was given a list of 22 emotion terms in alphabetical order; each term was accompanied by a brief descriptive scenario (Tables 1 and 3). Descriptive scenarios for each emotion were derived from conceptual analyses of the appraisals of each emotion (Lazarus, 1991; Oatley, Keltner, & Jenkins, 2006), and translated into terms more closely resembling our posers’ lay language of emotion. Posers were given time to contemplate the list of emotion terms and scenarios in anticipation of producing voiced expressions for each state without using words or word fragments. No explicit guidance was provided about which sorts of sounds posers should generate, nor were posers given examples to mimic. Posers were escorted into a private anechoic chamber (a soundproof room with no reverberation), outfitted with headphones, and positioned approximately 12 inches from a microphone. Posers then proceeded through the list of emotions and scenarios first indicating the emotion, then producing between one and five vocalizations for each emotion, at their own pace. The experimenter provided feedback throughout the vocal burst production process over a remote intercom, discouraging word or partial word use and inviting additional attempts.

Vocalizations were immediately digitized at 44 kHz and saved onto a computer hard drive. Vocalizations that included words or were inaudible were excluded, yielding 421 vocal burst stimuli: 15 to 20 conveying each emotional state. Mean length across all 421 vocal bursts of 22 different emotions was 1.55 s.

To make the duration of the judgment tasks tolerable and to make the lists of emotion term choices suitably complex, vocal burst stimuli were parsed into negative (anger, contempt, disgust, embarrassment, fear, guilt, sadness, shame, surprise; \( n = 172 \)) and positive (amusement, awe, compassion, contentment, desire, enthusiasm, gratitude, interest, love, pride, relief, sensory pleasure, triumph; \( n = 249 \)) stimulus sets before judging. Grouping vocal bursts into affect categories was also hypothesized to provide conservative measures of accuracy; vocal bursts within valence

\(^1\) Surprise can be thought of as a neutral valence emotion (Scherer et al., 2003). We include it with the negative emotions because in facial expression studies it is reliably confused with fear (Ekman et al., 1969).
Table 1
Negative Emotion Terms and Descriptive Scenarios Used to Prompt Vocal Bursts From Posers

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>You have been offended and intend to defend yourself</td>
</tr>
<tr>
<td>Disgust</td>
<td>You come in contact with something physically noxious or contaminating</td>
</tr>
<tr>
<td>Fear</td>
<td>You are faced with physical danger</td>
</tr>
<tr>
<td>Sadness</td>
<td>You experience the loss of a loved one</td>
</tr>
<tr>
<td>Surprise</td>
<td>You find out something unexpected that you didn’t know</td>
</tr>
<tr>
<td>Contempt</td>
<td>You feel superior to someone because of them doing something negative</td>
</tr>
<tr>
<td>Embarrassment</td>
<td>Someone discovers that you have made a social gaffe</td>
</tr>
<tr>
<td>Guilt</td>
<td>You know that you have done something (morally) wrong that has hurt another person</td>
</tr>
<tr>
<td>Shame</td>
<td>You feel like you are a bad person</td>
</tr>
</tbody>
</table>

Note. The exact terms and descriptive scenarios above were provided to posers in one alphabetized list that included both negative and positive terms. Negative emotion terms and descriptive scenarios, with pronouns and grammar changed to replace “You have . . . ” with “The person has . . . ”, were read to Study 1 and 3 judges prior to commencing the negative emotion and prototype vocal burst identification tasks.

Results

Across the 172 negative emotion vocal bursts, one-way analysis of variance (ANOVA) with poser sex as the independent variable and identification accuracy for each individually judged vocal burst as the dependent variable showed that judges (n = 26) were more accurate at identifying emotions conveyed by female than by male posers: 49% versus 42%, respectively, F(1, 4470) = 25, p < .001. Judges identified negative emotion vocal bursts from female posers an average of 12% more accurately, with the exception of...
Table 3
**Positive Emotion Terms and Descriptive Scenarios Used to Prompt Vocal Bursts From Posers**

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epistemological emotions</strong></td>
<td></td>
</tr>
<tr>
<td>Amusement</td>
<td>You find something humorous</td>
</tr>
<tr>
<td>Awe</td>
<td>You feel that you are in the presence of something greater than yourself</td>
</tr>
<tr>
<td>Interest</td>
<td>You are curious about something and want to explore it</td>
</tr>
<tr>
<td>Relief</td>
<td>Some unpleasant experience ceases</td>
</tr>
<tr>
<td><strong>Pro-social emotions</strong></td>
<td></td>
</tr>
<tr>
<td>Compassion</td>
<td>You are moved by someone else’s suffering</td>
</tr>
<tr>
<td>Gratitude</td>
<td>You feel appreciative of something given to you</td>
</tr>
<tr>
<td>Love</td>
<td>You are deeply affectionate towards someone</td>
</tr>
<tr>
<td><strong>Savoring emotions</strong></td>
<td></td>
</tr>
<tr>
<td>Contentment</td>
<td>You are deeply satisfied with your current condition</td>
</tr>
<tr>
<td>Desire</td>
<td>You are sexually attracted to someone</td>
</tr>
<tr>
<td>Sensory pleasure</td>
<td>You are enjoying something pleasing in taste or scent</td>
</tr>
<tr>
<td><strong>Agency-approach emotions</strong></td>
<td></td>
</tr>
<tr>
<td>Enthusiasm</td>
<td>You are looking forward to an event or experience</td>
</tr>
<tr>
<td>Pride</td>
<td>You have done something that increases people’s opinions of you</td>
</tr>
<tr>
<td>Triumph</td>
<td>You celebrate a victory</td>
</tr>
</tbody>
</table>

*Note.* The exact terms and descriptive scenarios above were provided to posers in one alphabetized list that included both positive and negative terms. Positive emotion terms and descriptive scenarios with pronouns and grammar changed to replace “You have . . .” with “The person has . . . ,” were read to Study 2 and 3 judges prior to commencing the positive emotion, and prototype vocal burst identification tasks.

Table 2 presents identification accuracy from judgments of the negative emotion vocal bursts. For all accuracy rates, binomial tests of significance were performed to ascertain the probability of obtaining accuracies observed, given chance levels of identification. Significance values are indicated by asterisk symbols in Table 2. Chance levels of accuracy were set in conservative fashion at 14.3 (1 out of 7), given the seven categories of emotion: anger, disgust, fear, sadness, surprise, and contempt, as well as self-conscious emotion (which include embarrassment, shame, and guilt) considered. This approach presupposes that accuracies for specific vocal bursts could be inflated by within-family guessing strategies, and has been used in similarly motivated studies of facial expression and touch (Frank & Stennet, 2001; Hertenstein, Keltner, App, Bulleit, & Jaskolka, 2006).

The well-studied emotions of anger, disgust, sadness, and surprise were well identified, replicating previous findings (Juslin & Laukka, 2003; Schroder, 2003). Though fear vocal bursts were identified with greater frequency than chance, they were more often labeled as surprise (37% vs. 46%, respectively). Vocal bursts for contempt, a less-studied negative emotion, were identified at above chance levels, although less consistently than the well-studied emotions. The ‘self-conscious’ emotions, embarrassment, guilt and shame, were identified at above chance levels, although only modestly so. As in studies of facial display (Haidt & Keltner, 1999), vocal bursts of ‘self-conscious’ emotions were often judged as sadness. It is also noteworthy that for contempt, as well as the ‘self-conscious’ emotion vocal bursts, a significant portion of judges used the ‘none of the above’ classification (>30%).

**Study 2: Judgments of Vocal Bursts for Four Categories of Positive Emotion**

**Method**

**Vocal burst stimuli.** Vocal bursts for the 13 positive emotions, collected as described in the method description for Study 1 (see Table 3), were used in Study 2.

**Judgment procedures.** The same judgment procedures that were carried out in Study 1 were carried out in Study 2, with the exception that the positive emotion vocal bursts and selection terms were presented. Two hundred forty-nine vocal bursts of 13 different positive emotions were presented to a second sample of 26 judges sex-matched to judges from Study 1 (20 female). After hearing each vocal burst judges selected the positive emotion term that best matched the emotion expressed. As in Study 1, Study 2 judges were presented with an alphabetized list of emotions terms matching the vocal burst categories of interest, as well as the option to select ‘none of the above.’

**Results**

Across all 249 vocal bursts of 13 positive emotions, a one-way AVOVA with poser gender as the independent variable and identification accuracy for each vocal burst judgment as the dependent variable confirmed that greater accuracy for female posers’ vocal bursts was significant: 36% versus 33%, respectively, $F(1, 6472) = 4.7, p < .05$. For vocal bursts of amusement, interest, relief, compassion, love, and enthusiasm, the advantage of female posers compared to male posers was 10%. For awe, contentment, desire, gratitude, and sensory pleasure vocal bursts, there was less than 2% difference in accuracy between female and male posed vocal bursts, and for pride and triumph, male posers’ vocal bursts were judged 12% more accurately than were those of female posers. As in Study 1, there were no differences in accuracy related to the sex of the judges.

Table 4 depicts the judgment accuracy for the 13 positive emotion vocal bursts evaluated in Study 2. As in Study 1, binomial tests of significance were performed, and significance indicated by asterisk symbols in Table 4. Chance was set at 25% (1 in 4) for the positive emotion vocal bursts to account for four families of positive emotion: ‘epistemological,’ ‘pro-social,’ ‘savoring,’ and ‘agency-approach,’ following the same conservative rationale described for determining chance levels in Study 1. Here one sees greater variability in the extent to which different positive states were communicated with vocal bursts. Judges proved to be quite adept at discriminating the vocal bursts of 3 of the 4 ‘epistemological’ emotions: amusement, interest, and relief. The vocal bursts of awe, compassion, sensory pleasure, enthusiasm, and to a lesser extent, triumph, were also identified at rates comparable to those
observed in studies of the face and voice (Elfenbein & Ambady, 2002). It was interesting that love and pride, known to have
detectable signals in face and postural behavior (Gonzaga et al.,
2001; Tracy & Robins, 2004), were not reliably detected in the
voice. Among the positive emotions, there were a few cases where
selection rates for nontarget emotion terms were significantly
greater than chance, that is, desire vocal bursts were frequently
judged to be sensory pleasure or interest (17% vs. 29% or 24%,
respectively) and triumph vocalizations were judged as enthusiasm
(29% vs. 42% respectively).

Study 3: Judgments of Prototypical Vocal Bursts for
Negative and Positive Emotion

Method

Vocal burst stimuli. For Study 3, a subset of the vocal burst
stimuli used in Studies 1 and 2 were selected for repeat judging. First,
all vocal bursts judged in Studies 1 and 2 were sorted according to
mean identification accuracy. Then, the four most accurately identi-
fied vocal bursts for each emotion were selected, with the constraint
that each of the four highly identifiable vocal burst were produced by
two different male and two different female posers. Thus, vocal bursts
with lower overall identification accuracy might have been selected
over a vocal burst with higher accuracy in order to preserve poser-
related heterogeneity. The four best-identified vocal bursts originating
from four different posers, two male and two female, for each emotion
were identified as our prototype vocal burst stimulus set. This proto-
type vocal burst set, including 36 negative, and 52 positive prototyp-
ical vocal bursts was used in Study 3.

Judgment procedures. Fifty-two new judges (39 female) each
performed two vocal burst identification tasks, one after the other.
Tasks presented either the prototype negative or the prototype positive
vocal bursts, with the same procedures and response options used in
Studies 1 and 2. Negative and positive prototype vocal burst judgment
tasks were presented in counterbalanced order across judges.

Results

Prototype vocal bursts of negative and positive emotions, there
were no differences in accuracy related to the sex of the posers or
the judges. Figure 1a depicts identification accuracy for the neg-
ative prototype vocal bursts judged in Study 3, alongside accuracy
rates for the full set of vocal burst stimuli judged in Study 1. Figure 1b depicts identification accuracy for the positive prototype
vocal bursts judged in Study 3, alongside accuracy rates for the full
set of vocal burst stimuli judged in Study 2. This figure illustrates
the breadth and boundaries of the voice’s capacity to communicate
different emotions.

In comparing accuracy for the full set of heterogeneous stimuli
(Studies 1 and 2) to accuracy for the prototypes (Study 3), it is
evident that some vocal bursts are expressed more uniformly
across posers (interest: no difference in accuracy between full set
and prototypes) while others were more variable between posers
(awe: sizable improvement in accuracy from full stimulus set to
prototypes). This may reflect the extent to which vocal bursts for
different emotions are closer to more universal, automatic ‘affect
bursts’ (amusement, anger, sadness, interest) or culturally in-
formed, paralinguistic ‘vocal emblems’ (awe, contempt).

Thus far, our analyses have examined the extent to which
observers could reliably judge vocal bursts of individual emotions.
A second interest was to examine whether a different picture might
emerge in the analyses of five emotion families: the ‘self-
conscious’ emotions, the ‘epistemological’ emotions, the ‘pro-
social’ emotions, the ‘savoring’ emotions, and the ‘agency-
approach’ emotions. Toward this aim, Table 5 depicts family-wise
accuracy rates for the negative and positive emotion vocal bursts
used in Studies 1, 2, and 3. Here, accuracy is defined as the use of
any of the related terms from an emotion family to judge a vocal
burst from that emotion family (e.g., selecting ‘embarrassment,’
’shame,’ or ‘guilt’ for any of the ‘self-conscious’ emotion vocal
bursts).

Table 4
Emotion Term Selection Rates for Vocal Bursts of 13 Positive Emotions

<table>
<thead>
<tr>
<th></th>
<th>Amusement</th>
<th>Awe</th>
<th>Interest</th>
<th>Relief</th>
<th>Compassion</th>
<th>Gratitude</th>
<th>Love</th>
<th>Contentment</th>
<th>Desire</th>
<th>Pleasure</th>
<th>Enthusiasm</th>
<th>Pride</th>
<th>Triumph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amusement</td>
<td>81*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awe</td>
<td>30*</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Interest</td>
<td>5</td>
<td>66*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Relief</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Compassion</td>
<td>24</td>
<td>13</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Gratitude</td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Love</td>
<td>6</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Contentment</td>
<td>8</td>
<td>7</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Desire</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Pleasure</td>
<td>10</td>
<td>11</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>29*</td>
</tr>
<tr>
<td>Enthusiasm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42*</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Pride</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Triumph</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29*</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>NOA</td>
<td>5</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>19</td>
<td>25</td>
<td>24</td>
<td>16</td>
<td>8</td>
<td>16</td>
<td>8</td>
<td>30</td>
<td>13</td>
</tr>
</tbody>
</table>

Note. The values represent mean percentages of emotion term selections for positive emotion vocal bursts; none of the above (NOA) rates are shown in
the bottom row. Blank cells correspond to selection rates of less than 5%. Data in bold indicate congruence with predictions.

*p < .05, chance levels set at 1 in 4, or 25%.
At this level of analysis, one sees a different degree to which emotions can be conveyed with vocal bursts. Judges were moderately accurate at identifying emotion from vocal bursts of individual emotions in the ‘self-conscious,’ ‘pro-social,’ ‘savoring,’ and ‘agency-approach’ families. However, at a family level, patterns of identification were more systematic. For example, judgments of prototype vocal bursts for compassion, gratitude, and love ranged from 5% to 44% correct. Yet 47% of the time, when presented with any of these ‘pro-social’ vocal bursts, judges identified them as one of these three ‘pro-social’ emotions. Similarly, identification of prototypical vocal bursts of embarrassment, guilt, and shame ranged from 20% to 29%; at the family level, they were correctly identified 47% of the time. These data lend credence to a family based approach to conceptualizing emotions.

**Discussion**

This investigation offers the most extensive examination to date of how many emotions the human voice can communicate. Vocal bursts were gathered from 10 untrained individuals, who were
given the task of conveying 22 different emotional states with their voice, based on theoretical descriptions of those states (Tables 1 and 3). In the first two studies, the full set of vocal bursts was presented to two groups of judges, who were asked to identify the emotions expressed by vocal bursts of negative (Study 1) or positive (Study 2) emotions. In Study 3, we presented the most well-identified vocal bursts from the first two studies, a prototype set, to new judges who were given the same forced-choice task of labeling the vocal bursts in emotion terms.

Identification of the well-studied, negative emotions—anger, disgust, fear, sadness and surprise—from vocal bursts ranged from 80% to 96% correct, corroborating existing findings (Juslin & Laukka, 2003; Sauter & Scott, 2007; Scherer, 1994). The three studies presented here provide the first evidence for vocal bursts of embarrassment, guilt, and shame, which were identified with low but above chance levels of accuracy, for vocal burst of awe and interest, which were readily identified, and for vocal bursts of compassion, gratitude, enthusiasm, and triumph, which were identified with low to moderate levels of accuracy. Taken together, these studies suggest that the voice can communicate at least 14 distinct emotional states without explicit words or obvious word substitutes.

With respect to the family-wise analysis of vocal burst emotion display, the present studies have several implications. First, although a family based approach to emotion has proven fruitful in understanding how individuals organize lexical terms for emotion (Shaver, Schwartz, Kirson, & O’Connor, 1987), the data described here are the first that we know of that suggest that emotion-related behaviors cluster into conceptually coherent families. This attests to the merits of family based approaches to studying emotions (Snowdon, 2003). These patterns of judgment begin to clarify the presence of food (savoring) or affiliative-caregiving coos and grins described in primates, such as specific calls that indicate the presence of food (savoring) or affiliative-caregiving coos and grins (pro-social), supporting the premise that humans have evolved vocalizations of emotion for comparable communicative purposes (Snowdon, 2003). These patterns of judgment begin to clarify the quality and resolution of the human voice’s capacity to communicate emotional information outside of speech.

There are methodological features of the present studies that lend confidence to our findings. The vocal bursts were provided by 10 randomly selected individuals, rather than trained actors or students of communication, as has been true of other studies of the face (Ekman, Sorenson, & Friesen, 1969) and voice (Scherer, 1986). In Studies 1 and 2, observers judged a large and heterogeneous sample of vocal bursts (rather than those preselected for their reliability of detection, as in Study 3). Accuracy rates, therefore, reflect a more general capacity for vocal bursts to express emotions, in spite of variability related to individual posers’ vocal

### Table 5

<table>
<thead>
<tr>
<th>Judgment study</th>
<th>Epistemological: change in understanding of the world</th>
<th>Pro-social: orienting towards others’ well-being</th>
<th>Savoring positive stimulation</th>
<th>Agency-approach: motivation to act towards improved prospects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies 1 and 2: n = 26, 172 negative/249 positive vocal bursts</td>
<td>35**</td>
<td>73**</td>
<td>30’</td>
<td>47**</td>
</tr>
<tr>
<td>Study 3: n = 52, 36 negative/52 positive prototypical vocal bursts</td>
<td>47”</td>
<td>87”</td>
<td>47”</td>
<td>66”</td>
</tr>
</tbody>
</table>

*Note.* The values represent mean percentages of family-wise emotion term selections for vocal bursts of ‘self-conscious’ (embarrassment, guilt, shame), ‘epistemological’ (amusement, awe, interest, relief), ‘pro-social’ (compassion, gratitude, love), ‘savoring’ (desire, contentment, sensory pleasure), and ‘status-enhancing’ (enthusiasm, pride, triumph).

"p < .05. ""p < .001.
qualities or expressive temperaments. Further, all three judgment studies incorporated the ‘none of the above’ response option, which reduces the likelihood that the accuracy rates obtained here were grossly affected by guessing strategies, which render forced choice judgment data problematic (Frank & Stennet, 2001). Finally, our criteria for statistical significance was set using conservative thresholds for chance accuracy, that is, taking emotion families into account, rather than simply basing chance on the exact number of emotions judged. These features increase confidence in the generalizability of the present findings.

At the same time, it is important to note important limitations of the present study, which point to future areas of inquiry. Our vocal bursts were posed, and not gathered during the experience of emotion (as is true of almost all emotion judgment studies). Thus, it will be important to ascertain whether similar levels of accuracy are obtained in judgments of vocal bursts that occur during real experiences of different emotions. While our samples of judges reflect the diversity of a large public university setting, they were all of a certain age and all lived in the United States. For claims of universality, it will be essential to collect and study the signal value of vocal bursts from individuals from different cultures (Sauter & Scott, 2007; Scherer et al., 2003). Furthermore, even with our ‘none of the above’ option, interpretation and generalization of accuracy rates presented here are limited by our use of a forced-choice task. A spontaneous recognition design could provide valuable supplemental information about single stimulus recognition accuracy in the domain of vocal bursts. The findings from the present study, limitations notwithstanding, reveal the voice to be perhaps the richest source of information about emotion, and new emotions such as awe and compassion, to be marked by distinct signals.

The voice is a prime modality for investigating emotion, capable of specific expression during and between speech incidents, with and without words. Here, we introduce how flexibly vocal bursts express specific states, informing basic understanding of emotion signals in social interaction and nonverbal communication.

References


Received March 11, 2008

Revision received December 16, 2008

Accepted July 23, 2009