Psychophysiology, 44 (2007), 827–837. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2007.00552.x

The effects of difficulty and gain versus loss on vocal physiology and acoustics

TOM JOHNSTONE,[a,b] CARIEN M. VAN REEKUM,[b] TANJA BÄNZIGER,[b] KATHRYN HIRD,[c] KIM KIRSNER,[a] AND KLAUS R. SCHERER[b]

[a]School of Psychology, University of Western Australia, Perth, Australia
[b]Psychology Department, University of Geneva, Geneva, Switzerland
[c]School of Psychology, Curtin University of Technology, Perth, Australia

Abstract

To examine the basis of emotional changes to the voice, physiological and electroglottal measures were combined with acoustic speech analysis of 30 men performing a computer task in which they lost or gained points under two levels of difficulty. Predictions of the main effects of difficulty and reward on the voice were not borne out by the data. Instead, vocal changes depended largely on interactions between gain versus loss and difficulty. The rate at which the vocal folds open and close (fundamental frequency; f_0) was higher for loss than for gain when difficulty was high, but not when difficulty was low. Electroglottal measures revealed that f_0 changes corresponded to shorter glottal open times for the loss conditions. Longer closed and shorter open phases were consistent with raised laryngeal tension in difficult loss conditions. Similarly, skin conductance indicated higher sympathetic arousal in loss than gain conditions, particularly when difficulty was high. The results provide evidence of the physiological basis of affective vocal responses, confirming the utility of measuring physiology and voice in the study of emotion.

Descriptors: Emotional prosody, Voice, Electroglottography, Physiology, Difficulty, Valence

The acoustic characteristics of natural, often involuntary markers of emotion in the voice and the physiological mechanisms responsible for such vocal modulation have as yet received little attention from researchers. In contrast to studies of acted emotional speech, the small number of studies on natural or induced emotional speech (e.g., Alpert, Kurtzberg, & Friedhoff, 1963; Bachorowski & Owren, 1995; Simonov & Frolov, 1973) have failed to identify acoustic patterns specific to different emotions. Instead, most acoustic changes measured in real or induced emotional speech can most parsimoniously be explained in terms of the general physiological arousal characteristic of different emotions. This contrasts with the theoretical view (e.g., Scherer, 1986) that emotions affect the acoustic characteristics of speech along multiple dimensions in a manner similar to the pattern of physiological changes that accompany emotion (e.g., Bradley, Codispoti, Cuthbert, & Lang, 2001; Cacioppo, Klein, Bernston, & Hatfield, 1993).

The reasons for the paucity of research on real or elicited emotional speech have been described elsewhere (e.g., Johnstone, van Reekum, & Scherer, 2001; Scherer, Johnstone, & Klasmeyer, 2003). These include the practical difficulty of inducing emotions in the laboratory and the fact that whatever emotional information is carried in speech is often shrouded or masked by other aspects (e.g., linguistic, social, cultural) of the speech signal, as well as the relatively recent development of affordable and effective speech analysis software and hardware. One of the major impediments has been the difficulty of obtaining high-quality, controlled sound recordings of real emotions. A solution to this has been the use of computer tasks and games to induce emotional responses, or at least components of emotional responses, in the laboratory (Johnstone, van Reekum, Hird, Kirsner, & Scherer, 2005; Kappas & Pecchinenda, 1999; MacDowell & Mandler, 1989). The use of such tasks also affords the ability to measure speech and physiology concurrently, with the potential for explaining measured acoustics of emotional speech in terms of the underlying physiological mechanisms. In this experiment, physiological measurements of autonomic nervous system (ANS) function were combined with electroglottographic (EGG; Fourcin & Abberton, 1971, 1977) measures of vocal fold function and acoustic analysis of speech while participants engaged in a computer task designed to evoke emotional responses.

The acoustic features that have most consistently been found in previous research to vary with different real or induced emotions are fundamental frequency (f_0), which is a measure of the frequency with which the vocal folds open and close, and speech energy, which depends on both the strength and efficiency of

Preliminary results from the study were presented at the New York Academy of Science meeting: “Emotions Inside Out: 130 Years after Darwin’s The Expression of the Emotions in Man and Animals,” New York, November 2002. This research was supported by Swiss National Science Fund grants 1114-037504.93 and 2151-049685.96.

Address reprint requests to: Tom Johnstone, Psychology Department, School of Psychology and Clinical Language Sciences, University of Reading, Whiteknights, PO Box 217, Reading, Berkshire, RG6 6AH, United Kingdom. E-mail: itjohnstone@gmail.com

vocal fold function as well as the resonance of the vocal tract. For high arousal emotions, such as anger, fear, or joy, f_0 and energy are commonly observed to be higher than for lower arousal affective states such as sadness or boredom (e.g., Alpert et al., 1963; Bachorowski & Owren, 1995; Duncan, Laver, & Jack, 1983; Johnstone et al., 2005; see Scherer et al., 2003). In a recent study that used a computer game to elicit emotional speech (Johnstone et al., 2005), f_0 and speech energy were both higher following game events that elicited negative emotions, such as losing a "life," compared to positive game events. These effects on f_0 and energy have been explained as resulting from an increase in the tension of the laryngeal muscles and an increase in subglottal air pressure that correspond to an elevation in sympathetic arousal (Scherer, 1986), although the empirical evidence for such an interpretation is lacking.

In this research, we anticipated that the use of EGG recording together with physiological and acoustic analysis would make the mechanisms underlying changes to f_0 and speech energy more clear. EGG is a technique that allows more direct measurement of changes to vocal fold dynamics in emotional speech than are afforded by acoustic speech analysis. With EGG, a small, high-frequency electric current is passed between two surface electrodes placed on either side of the speaker's neck, at the level of the larynx. Because the electrical impedance of the path between the two electrodes changes as the glottis opens and closes, a measurement of the impedance can be used as an indicator of glottal opening and closing (Fourcin & Abberton, 1971, 1977).

Of particular interest is the way changes to f_0 due to experimental manipulations would be reflected in the different parts of the glottal period (i.e., a single opening and closing cycle of the vocal folds). If sympathetic arousal is primarily responsible for f_0 changes through changes to laryngeal tension, we would expect f_0 variation across manipulations to correspond to similar variations in glottal phase measures as well as variation in ANS indicators such as skin conductance, finger temperature, and cardiac interbeat interval.

In addition to f_0 and energy changes to speech, a number of studies of acted emotional speech have also noted changes to spectral energy distribution across different expressed emotions (e.g., Banse & Scherer, 1996; Juslin & Laukka, 2001; Laukka, Juslin, & Bresin, 2005). Scherer (1986) predicted that constriction or opening of the vocal tract in negatively or positively valenced situations, respectively, would cause such spectral changes by selectively increasing or decreasing amplification of high frequency harmonics. Some supporting evidence for this hypothesis comes from the study of Tartter and Braun (1994), who noted an increase in second formant frequency when speakers were smiling. In the study of Johnstone et al. (2005), the relative proportion of energy in lower frequencies (below 1000 Hz) for voiced parts of speech recorded while playing a computer game varied in agreement with Scherer's predictions. Thus the proportion of energy at low frequencies (below 1000 Hz) was higher for pleasant than for unpleasant game situations.

It is possible that the spectral distribution of energy in affective speech reflects not only the configuration of the vocal tract but also vocal fold function linked to arousal. Research that has shown a link between vocal fold function and spectral slope (Fant, 1993) suggests that increased tension in the intrinsic laryngeal musculature causes vocal folds to remain closed over a longer portion of the glottal period, only briefly opening and then very rapidly closing again. The result is proportionally more energy in upper harmonics, which can result in speech with more high-frequency energy. Thus one might expect vocal expressions of emotions in which laryngeal muscles are more tense to be characterised by relatively greater energy in higher harmonics and a higher proportion of high-frequency energy. Using more direct measurements of vocal fold function, such as EGG, might make it possible to determine whether the spectral changes observed in emotional speech reflect glottal (i.e., vocal fold) or supraglottal (i.e., vocal tract) function.

In summary, we used EGG and peripheral psychophysiological measures in this experiment to provide evidence as to which mechanisms are responsible for affective changes to the f_0, energy, and spectral characteristics of speech. We used a computer task in which participants moved a small box on the screen either toward a moving target that gave them points (gain condition) or away from a target that subtracted points (loss condition). The controllability of the participant's box was adjusted to either make the task relatively easy or more difficult. Thus the effects of reward (gain vs. loss) and difficulty (low vs. high) on speech acoustics, vocal fold function, and ANS function were measured continuously while participants performed the task.

Predictions for the reward manipulation were based on Scherer's (1986) predictions of the acoustic consequences of the setting of the vocal tract and facial muscles in either an appetitive or aversive configuration as well as previous observations of valence effects on the spectral properties of speech (Johnstone et al., 2005). Thus for the condition in which participants were gaining points, we predicted an increase in the proportion of speech energy at low frequencies, reflecting an expansion and relaxation of the vocal tract. When participants were losing points, we predicted the opposite changes, namely, an increase in the proportion of high-frequency energy due to contraction and tensing of the vocal tract. As an independent measure of valence-specific changes to autonomic function, we measured finger temperature, which is affected by peripheral vasoconstriction or vasodilation. Peripheral vasoconstriction has been observed under conditions of threat, as opposed to challenge (e.g., Tomaka, Blascovich, Kelsey, & Leitten, 1993). We thus predicted skin temperature to be lower (corresponding to a greater degree of peripheral vasoconstriction) during loss as compared to gain situations. This prediction was also based on our past research, in which aversive emotional conditions in a computer game caused a drop in peripheral skin temperature (van Reekum et al., 2004).

With respect to difficulty, we predicted that increased task difficulty would result in heightened sympathetic arousal, as indicated by greater skin conductance activity and increased heart rate. Such arousal would be reflected in a corresponding increase in f_0. Furthermore, the manner in which the vocal folds opened and closed was predicted to change with task difficulty as a consequence of changes to laryngeal muscle tension and/or subglottal pressure. Specifically, high-difficulty situations, in which laryngeal muscles were predicted to be more tense, were expected to lead to relatively shorter open phase and longer closed phase.

As a consequence of changes to glottal phases in response to the difficulty manipulation, we also predicted a shallower spectral slope and a relatively lower proportion of energy at low frequencies in high-difficulty situations than in low-difficulty situations.

Finally, the two manipulations were expected to interact in their effect on physiology and the voice, such that the predicted task difficulty effects were expected to be greater in loss situations than in gain situations. This prediction is based on the notion that effectively coping with a challenging situation becomes even more relevant when the possible outcome of a failure to cope is negative (a corollary of the concept of a negativity bias; see Ito, Larsen, Smith, & Cacioppo, 1998; Kanouse & Hanson, 1987).

Method

This experiment was conducted as part of an ongoing, larger research project on the effects of emotion and stress on automatic speaker verification technologies, and made up part of a larger battery of emotion induction tasks. The total duration of the experimental session was 1.5 h, of which instructions, preparing the speaker for physiological and EGG measurements, and this experiment took approximately 45 min.

Speakers

Speakers were 30 male, French first-language adults recruited via written notices and local radio announcements. All speakers were paid SFr. 40 for their participation. Speakers were also told before the start of the session that if they performed well enough, they could win an extra SFr. 10. This incentive served to increase motivation and involvement in the task and thus render the manipulations more powerful. All speakers were, in fact, paid the extra SFr. 10 at the conclusion of the experimental session. Fully informed written consent was obtained from all speakers, who were free to withdraw from the experiment at any time. The research was carried out in compliance with University of Geneva human subjects research guidelines.

The Task

The task used in this experiment was a tracking task, in which the speaker had to use the mouse to control the movement of a small box on the screen. The task was presented in successive stages. In each stage, in addition to the player's box, one of two symbols, representing either gain or loss, respectively, moved about the screen. The loss symbol approached the player's box and the gain symbol avoided the player's box. The player's task was to keep the loss symbol from touching the box and to touch the gain symbol with the box. In addition to loss and gain, each type of symbol moved either quickly or slowly, thus making the task of avoiding the loss or achieving the gain more or less difficult. Thus gain versus loss, and difficulty, were manipulated in a factorial design.

Gain and loss were implemented using the addition and subtraction of points, respectively. The player's points were displayed in both a digital and graphical form to the left of the playing space. Points were continuously awarded or subtracted from the player's score on the basis of the distance of the player's box from the gain or loss symbol. The closer the player's box to the gain symbol, the more points he was awarded. The closer his box to the loss symbol, the more points he lost. All players started with 5000 points. Task difficulty was adjusted by varying the speed of the gain and loss symbols, so as to lead to a comparable performance across participants and within participants for the two easy conditions and the two difficult conditions, respectively. Note that because this adjustment was made individually for each participant and for each pair of conditions, it was not possible to objectively compare the absolute performance across participants or across conditions. The experiment consisted of four stages, one of each condition, presented in an order counterbalanced across subjects.

Speech Material

At four equally spaced intervals during each task stage, a small message at the bottom of the screen appeared, prompting the player to either pronounce a standard phrase (two times) or to pronounce an extended [a] vowel (two times). Sustained vowel production involves a highly regular cycle of vocal fold opening and closing that gives rise to a signal that can more precisely be measured and characterized than one that involves rapid and irregular transitions, as is the case with speech. In addition, speech production involves a much more complex, nonlinear interaction of vocal fold function with vocal tract dynamics and resonances. The extended vowel was thus chosen to allow maximum sensitivity to detect small affective changes to vocal function, as compared to the standard phrase, which would afford less sensitivity but would yield results more generalizable to natural speech. The order of standard phrase prompts and [a] prompts was randomized. The standard phrase was "Ceci est la tâche 4 2 5 1 0" ("This is task 4 2 5 1 0"), with a different five-digit combination of digits used in the phrase for each task stage. The digit combinations were counterbalanced across task stages and subjects. The correct way to respond to the two prompts, including instructions on how to pronounce the extended [a] vowel, was demonstrated to each participant prior to a practice stage. The four task stages then followed the practice stage.

Subjective Emotion Reports

Following every task stage, speakers were asked to report how they felt, using mouse-operated emotional state and intensity scales. With the scales, participants could choose any number of emotions from a given list of seven provided emotions. For each chosen emotion, they indicated an intensity ranging from *not felt at all* through *felt weakly*, *felt moderately*, to *felt extremely* on a continuous rating scale (stored internally as an x-pixel value ranging from 0 to 400). The emotions were satisfied, irritated, tired, stressed, disappointed, content, bored, and anxious. Alternatively, speakers could click a box indicating no felt emotion.

Measurements

Vocal measures. The acoustic speech signal was recorded using a Sennheiser clip-on condenser microphone (Model MKE 2 P-C, frequency response 20 Hz – 20 kHz $\pm$ 3 dB) connected to one channel of a Casio Digital Audio Tape (DAT) recorder. EGG electrodes were placed either side of the participant's larynx and secured with a Velcro band. The electrodes were connected to a Portable Laryngograph (Laryngograph Ltd.), which produces a high-frequency current of low voltage and amperage that passes through the glottis, with the amplitude varying as a function of impedance, which changes as the vocal fold contact area changes (for details, see Childers & Krishnamurthy, 1985). The Laryngograph output was connected to the other channel of the DAT recorder. These recordings were then transferred digitally to a PC and stored as 22-kHz sampling rate stereo wave files.

Physiological measures. All physiological measures were recorded continuously throughout the experimental part of the session with a Biopac MP100 physiology measurement system, at a sample rate of 250 Hz. Skin conductance was measured using 8 mm Ag-Ag/Cl electrodes placed on the tops of the index and middle fingers of the nondominant hand. The electrodes were filled with an NaCl paste (49.295 g of unibase and 50.705 g of isot. NaCl 0.9%). ECG was measured using pregelled, disposable ECG electrodes placed in a standard lead II (arm–leg) placement. Finger temperature was measured using a Biopac finger temperature probe attached to the small finger on the nondominant hand.

A 1-byte digital marker was output from the parallel port of the experimental presentation PC to the MP100 system. This channel marked the physiological data with the precise onset and offset of each task event, as well as the onset and offset of vocal recordings.

Procedure

On arrival, speakers were told that the purpose of the experiment was to collect a variety of speech recordings under different situations that were designed to simulate those that might occur during everyday use of computers. Speakers were informed that the program would prompt them at specified times to provide spoken input, which would be recorded.

The physiological sensors and EGG electrodes were then attached and the signals adjusted so as to be suitably amplified and free of noise. The microphone was then fitted and the recording level adjusted appropriately. Participants were told to move only the hand used for control of the mouse and to otherwise remain in a relaxed but still position. In particular, participants were asked to avoid moving their nondominant hand, which was placed in a comfortable resting position on the experimental desk. Participants were then asked if they were comfortable and ready to start before the experimenter left the experimental room and started the program.

At the end of the session, speakers were debriefed and reimbursed for their participation.

Data Reduction

EGG measures. Glottal open and closed times were calculated from the EGG signal using a combination of peak-picking and thresholding on the high-pass filtered (10 Hz cutoff), differentiated EGG signal (Henrich, d'Alessandro, Doval, & Castellengo, 2004; Marasek & Pützer, 1997). Instants of glottal closure, which are typically the points of maximum positive gradient in each glottal period, were identified as peaks in the differentiated EGG signal. Each glottal period, as demarcated by successive instances of glottal closure, was analyzed to identify open and closed times. An amplitude criterion of 90% of total period amplitude was used to determine the onset and offset of closed phase. The onset of glottal open phase was located using the equal level criterion, which is the point at which the glottal signal drops to a level equal to the level at the instant of glottal closure. It should also be noted that there are four phases in the glottal cycle: opening, open, closing, and closed. In this experiment we did not measure the opening and closing phases because of a lack of reliability in the measurements due to their brief, transient properties.

In addition to the measurement of *absolute* open and closed times, glottal quotients, which indicate the *proportion* of each glottal period occupied by the open and closed phases, were also measured. In contrast to absolute times, quotients measure the shape of the glottal waveform after normalization to the length of the glottal period and are thus independent of the absolute f_0 level. Low-frequency (below 10 Hz) RMS EGG energy was also examined as an exploratory potential measure of larynx movement associated with intonation during pronunciation of the standard phrase.

Acoustic measures. The set of acoustic measures was chosen on the basis of which measures have been shown to vary with different expressed emotions in previous studies (see Scherer et al., 2003). A slightly different selection of vocal measures was chosen for the extended vowel and the standard phrase, because the dynamics of the two types of vocalization are distinct. This is explained in the description of measures provided below. All acoustic analyses were performed using in-house software running under LabVIEW (National Instruments, Inc.) and based on algorithms in Deller, Proakis, and Hansen (1993).

Fundamental frequency. f_0 was estimated from successive glottal closure instants in the EGG signal. Median f_0, rather than mean f_0, was analyzed in this experiment, because median f_0 is less susceptible to the effects of outliers and the skewed distribution of f_0 values. The standard deviation of f_0 values has previously been used as a measure of overall f_0 variability (e.g., Banse & Scherer, 1996). However, it is difficult to know whether the standard deviation reflects global, suprasegmental variation in f_0 or period-to-period variation. The former corresponds to the changes in f_0 associated with intonation, whether emotional or non-emotional, and is related to the f_0 range (difference between f_0 floor and ceiling). The latter type of variation, termed jitter, has been independently measured as an indicator of emotional stress in previous research (e.g., Smith, 1977). In this experiment, measurement of the EGG signal made possible the accurate period-to-period measurement of f_0, so that jitter could be directly quantified for the extended vowel vocalization (for the standard phrase, large intonational deviations of f_0 make jitter a nonreliable measure). To quantify jitter, a quadratic curve was fit to a running window of five successive f_0 values using a least mean squares curve-fitting algorithm. The quadratic curve was then subtracted from that section of the f_0 contour. This served to remove longer term f_0 movements, which would otherwise contaminate jitter measurements. Jitter was then calculated as the mean of the magnitude of period-to-period variation in the residual f_0 values. As a measure of intonational f_0 range during the standard phrase, we measured the f_0 floor and f_0 ceiling, quantified as the 5% percentile and 95% percentile, respectively, of the distribution of f_0 values for each separate utterance (see Banse & Scherer, 1996).

Speech energy. Mean RMS energy was calculated from voiced segments of the acoustic speech signal.

Spectral slope. The power spectrum of voiced speech has a spectral slope of approximately -6 dB per octave. The source of this spectral slope is the spectrum of the glottal excitation, which has a slope of approximately -12 dB, which is then modified by the filtering of the vocal tract and lips (see Deller et al., 1993; Fant, 1960). Although the glottal source spectrum is not directly available for measurement, it is possible that the spectral slope of the acoustic speech signal will vary with emotion-induced changes to the glottal source spectrum (Klasmeyer & Sendlmeier, 1997; Scherer, 1986). To characterize spectral slope, a regression line was fit to the voiced average power spectrum of the acoustic signal and the slope of the line in dB/octave was calculated.

Physiological measures. The skin conductance signal was low-pass filtered, with a cutoff of 0.7 Hz, and down-sampled to 20 Hz before parameter extraction. An automated routine enabled the scoring of the number of skin conductance responses as well as the skin conductance response amplitudes. A skin conductance response was scored when the increase in skin conductance level exceeded 0.05 μS. The amplitude of the maximum skin conductance response during each analysis window was used as a measure of the peak autonomic response during that experimental period. In addition, the mean skin conductance level for segments not corresponding to skin conductance responses was determined. All three measures are thought to relate slightly differently to the underlying autonomic physiology and have been previously employed in somewhat different experimental contexts (e.g., Dawson, Schell, & Filion, 1990; Nikula, 1991; Öhman, Esteves, Flykt, & Soares, 1993; van Reekum et al., 2004).

The cardiac interbeat interval (IBI) derived from the R-R interval was calculated using an automatic peak picking routine. Those IBIs shorter than 400 ms and longer than 1500 ms were considered as artifacts and eliminated from further analysis. Both mean IBI and standard deviation of IBI (as a measure of heart rate variability) were derived. Physiological measures were only scored during periods of no speech in the 10-s window preceding each vocal report, so as to avoid speech-related contamination of the physiological signals.

Statistical Analysis

Because of the large number of dependent measures, we adopted an analysis strategy that would guard against alpha inflation but still permit an examination of individual acoustic, EGG, and physiological parameters. In a first step, multiple dependent measures were reduced to a small number of factors using Principal Components Analysis (PCA) with direct Oblimin rotation, estimated on the basis of measures that had been residualized with respect to subject. The residuals were used because the factor structure we were interested in was that due to variation within subjects across experimental conditions, and not variation across subjects that could have many nonpsychological causes (e.g., skin thickness in the case of skin temperature and skin conductance measures). The Kaiser criterion (Kaiser, 1960) was used to determine the number of factors to retain for further analysis (i.e., factors with eigenvalues $\geq$ 1 were retained). In the second step, these factors were analyzed with a mixed effects GLM with reward and difficulty as fixed factors and subjects as a random factor. One such analysis was carried out for the physiological data, and two separate analyses were performed for the vocal data corresponding to the recorded standard phrase and the extended vowel, respectively. Separate analyses were performed because the relationship between variables was expected to differ for standard phrase and extended vowel recordings and because a slightly different selection of vocal measures was analyzed for the extended vowel and the standard phrase.

Where the GLM tests on a given factor score indicated significant effects, we followed up with univariate tests on the individual measures that loaded on that factor to identify those measures that contributed to the effect.

Results

Emotion Reports

Figure 1 shows the mean reported intensity of contentment, disappointment, stress, and satisfaction for each condition. These four emotions were the only ones reported consistently across subjects. The four rated emotions were analyzed with reward by difficulty by speaker mixed model ANOVA. Only rated intensity of stress and satisfaction differed as a function of the manipulations. Stress was higher for difficult than for easy conditions, $F(1,28) = 4.7$, $p = .04$. Satisfaction was higher for easy than for difficult conditions, $F(1,28) = 10.9$, $p = .003$.

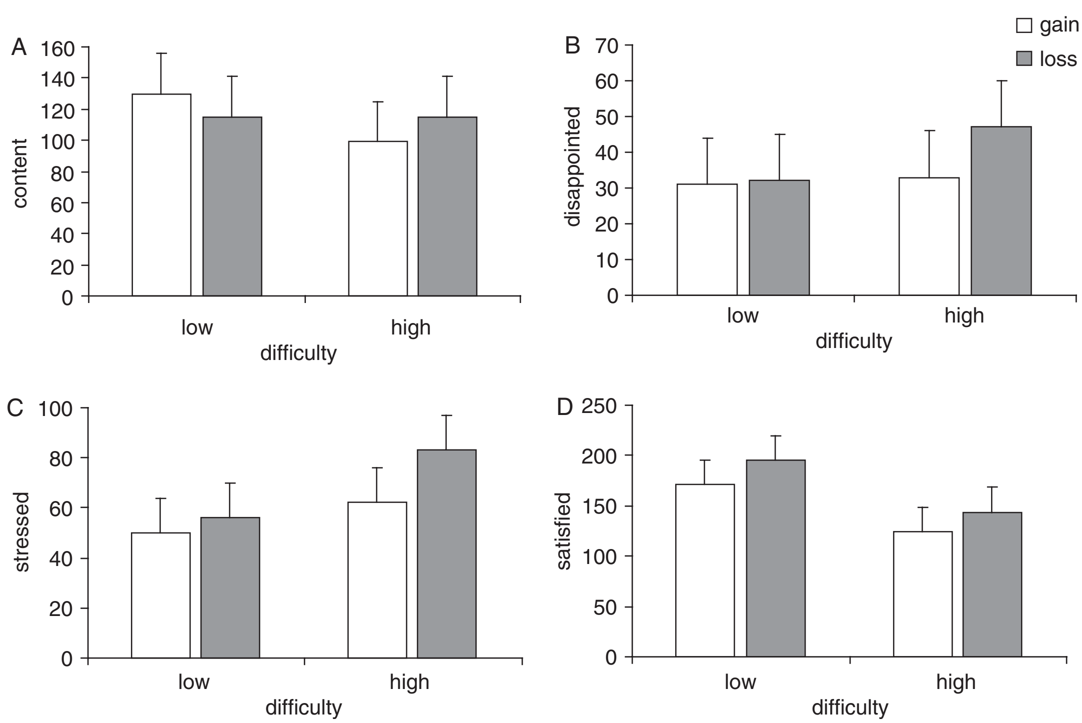

Figure 1. Subjective ratings of felt contentment (A), disappointment (B), stress (C), and satisfaction (D) as a function of valence and difficulty. Bars represent 95% within-subject confidence intervals. Note that the scales of the y-axis are different for the different emotions.

Acoustic and EGG Measures for the Extended [a] Vowel

For the extended [a] vowel, PCA analysis yielded three factors that explained 74% of the variance (see Table 1, top). The first factor, which accounted for 41% of the variance, had high positive loadings of closed time and quotient and negative loading of open time and quotient. The second factor, accounting for 20% of the variance, had high positive loadings of median f_0, speech energy, and spectral slope and negative loading of open time and spectral energy under 1000 Hz. The third factor was characterized by a positive loading of jitter and a negative loading of spectral energy under 1000 Hz.

GLM analysis of the first factor yielded a main effect of reward, $F(1,28) = 5.65$, $p = .02$, but no effect of difficulty, $F(1,28) = 0.32$, $p = .57$, nor any interaction of Difficulty × Reward, $F(1,27) = 1.88$, $p = .18$. Examining the vocal measures that loaded on the first factor, closed quotient was higher in loss than in gain conditions, $F(1,28) = 5.64$, $p = .03$, although closed time was not significantly different, $F(1,28) = 2.86$, $p = .10$. Both open time, $F(1,28) = 7.16$, $p = .01$, and open quotient, $F(1,28) = 6.66$, $p = .02$, were lower in loss than in gain conditions.

The second factor showed a significant effect of reward, $F(1,28) = 5.90$, $p = .02$, and a marginal interaction of Difficulty × Reward, $F(1,27) = 3.22$, $p = .08$. No main effect of difficulty was apparent, $F(1,28) = 0.22$, $p = .64$. Of the measures loading on this factor, median f_0 was marginally higher for loss than for gain conditions, $F(1,28) = 3.59$, $p = .07$, but also showed a significant interaction of Difficulty × Reward, $F(1,27) = 5.79$, $p = .02$, indicating that this difference was significant for the high-difficulty condition ($p = .02$) but not for the low difficulty condition ($p = .84$; though note that the interaction for this factor was only marginally significant). No significant effects involving the reward manipulation were found for spectral slope, reward: $F(1,28) = 1.920$, $p = .18$; Reward × Difficulty: $F(1,27) = 0.15$, $p = .70$, speech energy, reward: $F(1,28) = 2.14$, $p = .16$; Reward × Difficulty: $F(1,27) = 0.97$, $p = .33$, or spectral energy under 1000 Hz, reward: $F(1,28) = 2.99$, $p = .10$; Reward × Difficulty: $F(1,27) = 1.23$, $p = .28$.

No significant effects were observed for the third factor, difficulty: $F(1,28) = 0.06$, $p = .82$; reward: $F(1,28) = 0.14$, $p = .71$; Reward × Difficulty: $F(1,27) = 0.02$, $p = .88$.

Means and confidence intervals for the acoustic and EGG measures of the extended [a] vowel that changed as a function of difficulty and reward are shown in Figure 2. Means and confidence intervals for all other measures are provided in Table 2.

Table 1. *Structure Matrices for the Three Principal Component Analyses (PCA) Performed, Showing Factor Loadings for the Different Vocal and Physiological Measures*

	Component Loadings			
	1	2	3	4
[a] vowel vocal measures				
Median f_0	.352	.713	.141	
Voiced energy	.187	.794	.136	
Jitter	−.119	−.054	.897	
Energy below 1000 Hz	−.141	−.553	−.508	
Spectral slope	−.018	.688	−.251	
EGG closed time	.850	−.143	−.046	
EGG closed quotient	.955	.299	−.004	
EGG open time	−.898	−.460	−.061	
EGG open quotient	−.911	−.174	.016	
Standard phrase vocal measures				
Median f_0	.100	.909	.288	−.050
Ceiling f_0	.085	.756	.151	−.155
Floor f_0	−.076	.728	.200	.247
Voiced energy	.085	.273	.839	.099
Energy below 1000 Hz	.082	−.179	−.820	.054
Spectral slope	.113	.008	−.129	.846
LF EGG energy	.083	−.016	.476	.630
EGG closed time	.765	−.543	−.107	.050
EGG closed quotient	.932	.094	.141	.111
EGG open time	−.786	−.569	−.092	−.104
EGG open quotient	−.932	−.012	.104	−.096
Physiological measures				
Interbeat interval (IBI)	−.651	.488		
IBI variability	−.055	.676		
Finger temperature	−.219	−.720		
Skin conductance level	.824	.314		
# Skin conductance responses	.554	−.084		
Skin conductance amplitude	.777	.164		

Acoustic and EGG Measures for the Standard Phrase

PCA analysis of the standard phrase vocal measures yielded four factors that accounted for 77% of the variance (see Table 1, middle). The first factor, accounting for 28% of the variance, had positive loadings of closed time and quotient and negative loadings of open time and quotient. The second factor, accounting for 25% of the variance, was characterized by positive loadings of the f_0 measures as well as a negative loading of closed time and open time. The third factor, accounting for 13% of the variance, had a positive loading of speech energy and a negative loading of spectral energy under 1000 Hz. The fourth factor had positive loadings of spectral slope and low-frequency EGG energy.

For the first factor, there were no significant effects of difficulty, $F(1,28) = 0.86$, $p = .36$, reward, $F(1,28) = 1.25$, $p = .27$, nor any interaction of Difficulty × Reward, $F(1,27) = 0.50$, $p = .49$.

Analysis of the second factor revealed a significant main effect of difficulty, $F(1,28) = 4.10$, $p = .05$, no effect of reward, $F(1,28) = 2.18$, $p = .15$, but a significant Difficulty × Reward interaction, $F(1,27) = 6.07$, $p = .02$. Median f_0 showed a main effect of difficulty, $F(1,28) = 4.22$, $p = .049$, but no interaction of reward with difficulty, $F(1,27) = 1.62$, $p = .21$. An interaction of Difficulty × Reward on f_0 ceiling, $F(1,27) = 9.26$, $p = .005$, was due to higher f_0 ceiling for loss than for gain under high-difficulty ($p = .01$) but not under easy conditions ($p = .52$). An interaction of Difficulty × Reward for f_0 floor, $F(1,27) = 4.92$, $p = .04$, was due to f_0 being marginally higher in difficult than in easy conditions when faced with loss ($p = .06$), but no difference when faced with gain ($p = .89$). These changes to f_0 were reflected in a similar interaction effect on glottal open time, $F(1,27) = 4.33$, $p = .05$, which was due to shorter glottal open time in difficult than in easy conditions when faced with loss ($p = .02$), but no difference when faced with gain ($p = .73$). Glottal closed time showed no such interaction effect, $F(1,27) = 0.40$, $p = .53$, nor any effects of reward, $F(1,28) = 0.02$, $p = .89$, or difficulty, $F(1,28) = 0.64$, $p = .43$.

The third factor showed no effect of difficulty, $F(1,28) = 1.70$, $p = .20$, reward, $F(1,28) = 1.04$, $p = .32$, nor Difficulty × Reward, $F(1,27) = 0.87$, $p = .36$.

The fourth factor showed no effects of reward, $F(1,28) = 0.19$, $p = .67$, or difficulty, $F(1,28) = 0.26$, $p = .61$, but did show a significant Difficulty × Reward interaction, $F(1,27) = 11.9$, $p = .002$. Driving this effect was an interaction of Difficulty × Reward on spectral slope, $F(1,27) = 11.75$, $p = .002$, due to spectral slope being shallower (i.e., less negative) for loss than for gain conditions under high difficulty

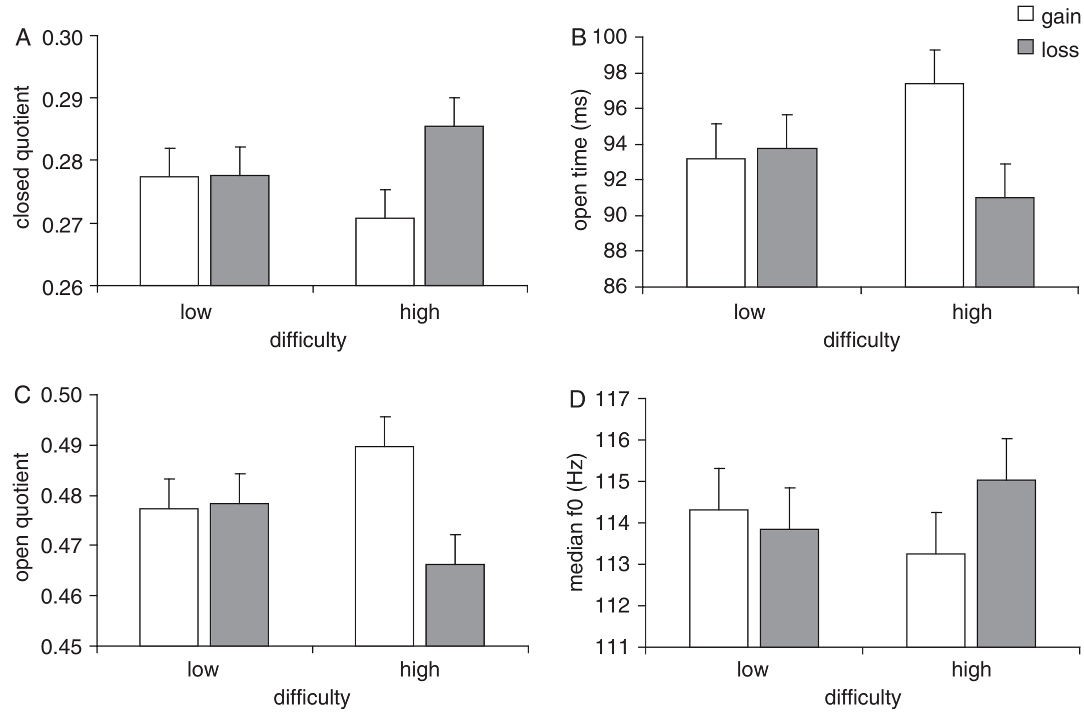

Figure 2. Acoustic and EGG measures for the extended [a] vowel that showed changes as a function of reward and difficulty. A: EGG closed quotient, B: EGG open time, C: EGG open quotient, D: median f_0. Bars represent 95% within-subject confidence intervals.

($p = .001$) but not under easy conditions ($p = .09$). Low-frequency EGG energy showed an interaction of Reward × Difficulty, $F(1,27) = 4.37$, $p = .05$, which was due to more low-frequency EGG energy for loss than for gain conditions under high-difficulty ($p = .04$) but not under easy conditions ($p = .16$).

Means and confidence intervals for the vocal measures of the standard phrase that changed as a function of difficulty and reward are shown in Figure 3. Means and confidence intervals for all other measures are provided in Table 2.

Table 2. *Mean values for the vocal and physiological measures as a function of difficulty and reward. Measures already shown in Figures 2–4 are not included. CI: 95% within-subjects confidence interval. The energy measures are given in analog to digital units.*

Difficulty	Low		High		
Reward	Gain	Loss	Gain	Loss	CI
Vocal measures: [a] vowel					
RMS energy	4727	4915	4569	4905	456
Jitter (Hz.)	0.568	0.579	0.722	0.538	0.369
closed time (ms)	54.33	54.59	53.28	55.60	1.71
energy < 1000 Hz (%)	87.18	86.78	87.15	86.64	0.012
spectral slope (dB/octave)	− 8.42	− 8.32	− 8.41	− 8.30	0.23
Vocal measures: phrase					
RMS energy	3346	3345	3369	3385	169
closed time (ms)	51.25	51.39	51.19	51.73	1.22
closed quotient	0.287	0.285	0.287	0.291	0.006
open time (ms)	84.70	85.15	84.56	82.32	1.76
open quotient	0.469	0.469	0.469	0.458	0.007
energy < 1000 Hz (%)	94.35	94.78	94.33	94.43	0.01
Physiological measures					
IBI (ms)	729	723	730	725	11.6
IBI variability (ms)	47.6	46.7	42.5	42.8	6.0
finger temperature (F)	87.95	87.76	88.09	87.73	0.13

Physiological Measures

PCA of the physiological variables yielded two factors that explained 57% of the variance (see Table 1, bottom). The first factor, which explained 35% of the variance, was characterized by high positive loadings of number of skin conductance responses, skin conductance response amplitude, and tonic skin conductance level and negative loading of interbeat interval. The second factor, explaining 22% of the variance, had a high negative loading of finger temperature, a positive loading of interbeat interval variability, and also a smaller positive loading from interbeat interval.

GLM analysis of the first factor showed a significant main effect of reward, $F(1,25) = 8.03$, $p = .01$. There was no main effect of difficulty, $F(1,25) = 0.00$, nor any interaction of Difficulty × Reward, $F(1,23) = 1.73$, $p = .20$, for the first factor. Examining the physiological measures that loaded on the first factor, tonic skin conductance level was higher for loss than for gain situations, $F(1,25) = 6.30$, $p = .02$, as were maximum skin conductance response amplitudes, $F(1,24) = 4.8$, $p = .04$. The number of skin conductance responses did not show a main effect of gain versus loss, $F(1,25) = 1.69$, $p = .21$. Cardiac interbeat interval was not significantly different between gain and loss conditions, $F(1,25) = 2.04$, $p = .17$.

Analysis of the second factor showed no significant effects of difficulty, $F(1,25) = 0.47$, $p = .50$, nor reward, $F(1,25) = 0.03$, $p = .87$, nor a significant interaction of Difficulty × Reward, $F(1,23) = 2.81$, $p = .11$.

Means and confidence intervals for the physiological measures that changed as a function of difficulty and reward are

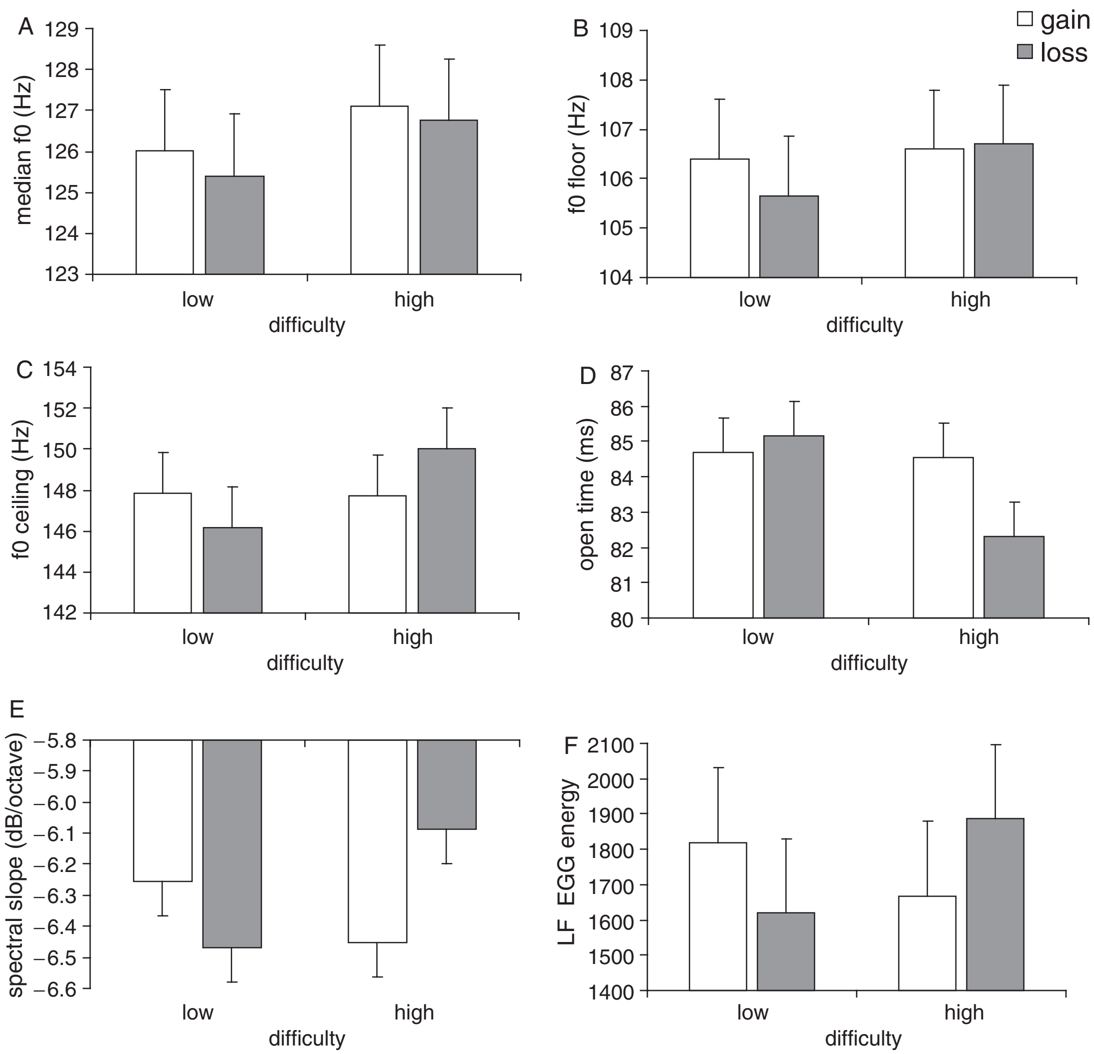

Figure 3. Acoustic and EGG measures for the standard phrase that showed changes as a function of reward and difficulty. A: median f_0, B: f_0 floor, C: f_0 ceiling, D: glottal open time, E: spectral slope, F: low frequency EGG energy. Bars represent 95% within-subject confidence intervals.

shown in Figure 4. Means and confidence intervals for all other physiological measures are provided in Table 2.

Discussion

Efficacy of the Manipulations

Confirmation of the efficacy of the experimental manipulations comes from both subjective reports and physiological data. Subjective report data reflected the difficulty manipulation, with players reporting feeling more stressed in difficult than in easy conditions and more satisfied in easy than in difficult conditions. There were no effects of gain versus loss on subjective reports. The skin conductance level and response amplitude data, however, indicate that players were more sympathetically aroused in loss conditions than in gain conditions. Thus, although subjective reports indicated no change in the players' subjective feelings in response to the gain and loss conditions, physiological and vocal measures (see below) did show that the manipulations produced measurable effects.

The manipulation of gain versus loss in this game was necessarily fairly subtle, for both ethical and practical reasons. It is thus possible that effects on subjective feelings were too small to be measurable using standard subjective reports. Vocal and physiological measures in such circumstances might provide a more sensitive indicator of emotional responses. These results thus highlight the way in which subjective feeling reports and physiological measures provide complementary information about emotion responses in different types of experimental situations and strengthen the case for routinely recording physiology in addition to subjective reports during experiments on emotion.

Main Effects of Reward

Based upon the predictions of Scherer (1986) and a previous study (Johnstone et al., 2005), we predicted that when participants were gaining points, there would be an increase in the proportion of speech energy at low frequencies, reflecting an expansion and relaxation of the vocal tract, relative to when participants were losing points and the vocal tract was more tense. As a physiological response to the reward manipulation, we predicted that finger temperature would be lower for the loss condition than for the gain condition (based on the finding of

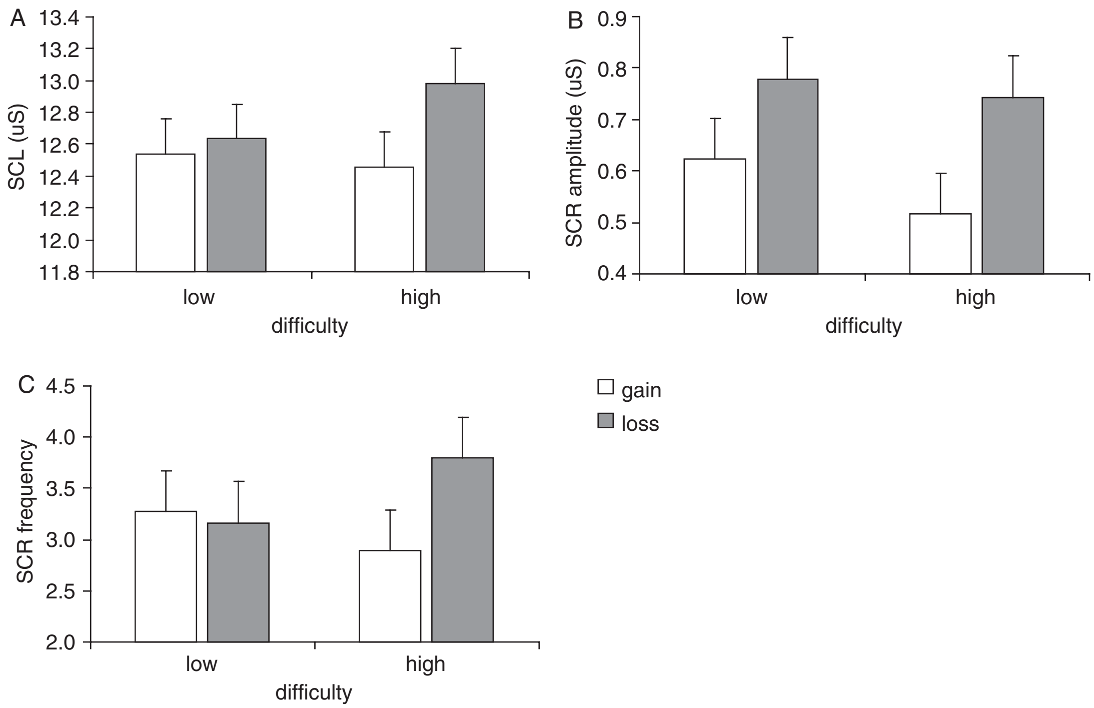

Figure 4. Physiological measures that showed changes as a function of reward and difficulty. A: tonic skin conductance level, B: maximum skin conductance response amplitude, C: number of skin conductance responses. Bars represent 95% within-subject confidence intervals.

Tomaka et al., 1993, of greater peripheral vasoconstriction under threat than under challenge). Neither of these predictions was supported by the data, with spectral energy under 1000 Hz, spectral slope, and finger temperature showing no main effect of the reward manipulation. Although both measures of spectral energy distribution have shown effects in previous studies (e.g., Banse & Scherer, 1996; Johnstone et al., 2005), neither are very sensitive measures. It might be necessary to include a more precise measurement of formant bandwidth and energy in future studies to address predictions of spectral energy change under conditions of gain or loss.

Although not predicted, we did discover differences between gain and loss conditions in skin conductance level and response amplitude, however, which suggests that the reward manipulation did in fact produce an autonomic response, with greater sympathetic activity in loss situations than in gain situations. These differences were paralleled by changes to the glottal cycle for the extended [a] vowel, with glottal open quotient larger for gain than for loss and closing quotient larger for loss than for gain. Thus under conditions of losing relative to gaining points, players showed greater sympathetic nervous system activity and the vocal folds remained closed for a greater proportion of the glottal cycle, at the "expense" of the period during which the folds are open.

Such an effect on glottal cycle corresponds to the description of the effects of increased laryngeal muscle tension given by Sundberg (1995), according to which the vocal folds are held in a closed position with greater force. A higher subglottal pressure is therefore required to build up so as to force open the vocal folds. When the vocal folds do eventually open, the rush of air through the glottis causes a sudden drop in pressure due to the Bernoulli effect. The vocal folds are thus rapidly drawn back together again, both under the influence of elastic forces of the increased muscle tension and the pressure drop between them. The result, termed "tense voice," is a more skewed glottal pulse, with longer closed phase and shorter open phase. Such was the shape of the glottal pulse with the [a] vowel under loss compared to gain situations in this experiment. Tense voice might thus be a feature of vowel production when in a sympathetically aroused state. No such effect was found for the standard phrase, which could be due to underlying differences in the way physiology affects the production of fluent speech as compared to extended vowels. Alternatively, it is possible that the dynamic aspects of f_0 control subserving intonation (possibly indicated by the Difficulty $\times$ Reward interaction observed for low-frequency EGG energy, which we measured as an exploratory measure of larynx movement) when pronouncing the standard phrase swamped the small emotion-mediated changes to glottal cycle. Given the general lack of effects for glottal measures on the standard phrase, it is probable that glottal parameters have limited sensitivity when averaged across multiple different vowels and across sentences where intonation causes large shifts to f_0.

Main Effects of Difficulty

For the difficulty manipulation, we predicted that increased task difficulty would result in heightened sympathetic arousal, as indicated by greater skin conductance activity and increased heart rate. Such arousal would be reflected in a corresponding increase in f_0. Furthermore, the manner in which the vocal folds opened and closed was predicted to change with task difficulty as a consequence of changes to laryngeal muscle tension and/or subglottal pressure. Specifically, high difficulty situations, in which laryngeal muscles were predicted to be more tense, were expected to lead to relatively shorter open phase and longer closed phase. As a consequence of changes to glottal phases in response to the difficulty manipulation, we also predicted a shallower spectral slope, and a relatively lower proportion of energy at low frequencies, in high-difficulty situations than in low-difficulty situations.

Only median f_0 for the standard phrase showed the predicted effect of difficulty, with a higher median f_0 for difficult than for easy conditions. Contrary to our predictions, there were no main effects of the difficulty manipulation on either the physiological measures or any other vocal measures.

Interaction of Difficulty and Reward

We had predicted an interaction of difficulty and reward, such that effects of task difficulty would be greater in loss situations than in gain situations. In fact, in this experiment we observed interactions for a number of vocal measures such that effects of the reward manipulation were greater in difficult than in easy situations. For the extended [a] vowel, glottal open time was shorter and median f_0 was higher for loss than for gain in difficult conditions but not for easy conditions. For the standard phrase, higher f_0 ceiling (indicating greater f_0 range) was measured for loss than for gain under difficult but not under easy conditions. Of the glottal parameters measured from the standard phrase, only glottal open time showed an interaction of difficulty and reward, reflecting shorter opening time for difficult than for easy conditions, when faced with loss, a result that paralleled the shorter f_0 floor for difficult loss than for easy loss conditions.

For the standard phrase, an interaction effect of difficulty and reward on spectral slope was also measured. Spectral slope was flatter for difficult loss than for difficult gain conditions. This result raises the question of whether the changes observed to spectral slope were due to changes in vocal fold function. The lack of corresponding changes to glottal quotient measures and the lack of a corresponding spectral slope effect for the extended [a] vowel, for which glottal effects were observed, would seem to suggest that no such clear link exists.

In summary, although reward and difficulty were manipulated orthogonally in this study, the two dimensions were found to interact in their effect on vocal production. Contrary to the predictions of Scherer (1986) and our hypotheses, the effects of the difficulty manipulation on vocal measures in this study were almost exclusively in the context of a possible loss situation. One possible explanation for the lack of a main effect of difficulty is that vocal responses to emotional situations are organized around the combined significance and meaning of a given situation, rather than separately according to its constituent properties. Many theories of emotion advance such a view. Scherer (1986) holds that it is the cumulative combination of evaluations of a situation along different dimensions that gives rise to an emotion response. Lazarus (1991) holds that the total evaluated meaning of a situation gives rise to a coordinated adaptive response, while Frijda (1986) posits that emotions correspond to adaptive action tendencies. Such coordinated mechanisms do not preclude the possibility that, under situations different from those examined in this study, single dimensions such as valence or difficulty might have distinct measurable effects on the voice. However, at least in the case of the vocal measures examined herein, it seems that the effect of a change in one emotion dimension often depends on other dimensions.

It is worth noting that the physiological measurements made prior to recorded speech in this experiment were necessarily indirect measures of vocal physiology. Although in principle it is possible to directly measure laryngeal muscle tension, lung volume, subglottal pressure, and the position and tension of vocal tract muscles, such measurements are extremely invasive and would thus interfere with the emotion induction procedure. Nevertheless, EGG recordings were successfully used to "bridge the gap" between acoustic and physiological measures when using a suitably controlled vocal utterance such as the extended vowel. The general method could potentially be extended to more varied types of vocalization, including connected speech, if enough tokens of each vowel in a similar linguistic and phonetic context were collected to allow a segment-by-segment analysis.

Conclusions

This research has differed from most previous research on emotional speech in two aspects. The emotional speech examined was not acted, but rather elicited using manipulations of a computer task along two primary emotion-relevant dimensions. Furthermore, acoustic and electroglottographic recordings of standard phrases and vowels were made in conjunction with physiological measurements in an effort to elucidate some of the mechanisms responsible for the resulting emotional changes to speech.

The results of this experiment suggest that, faced with a situation entailing loss, the body mobilizes its resources to actively cope by elevating activity in the sympathetic branch of the autonomic nervous system, as indicated by skin conductance level and skin conductance response amplitude. This elevated sympathetic activity corresponds to a probable increase in laryngeal tension, which leads to a corresponding change in the opening and closing of the vocal folds. In addition to a general speeding up of the vocal cycle, the folds close with more force and remain closed for a longer proportion of each vocal cycle (Sundberg, 1995).

The combined use of acoustic, EGG, and physiological measurements in this experiment has provided valuable insights into the changes that occur in vocal fold function due to emotional physiological responses. Measured changes, though significant, were small, however, which, in conjunction with the lack of subjectively felt emotion, points to the relatively mild emotional responses elicited. Perhaps as a consequence, the study has not provided clear results with respect to spectral changes that have been observed in previous research and has more limited sensitivity to the analysis of fluent speech as compared to extended vowel production. Further use of EGG and physiological measurements coupled with more sensitive segmental and formant analysis will yield not only better understanding of emotional vocal production but also add to our knowledge of how emotional responses in various expressive and physiological modalities are organized.

REFERENCES

Alpert, M., Kurtzberg, R. L., & Friedhoff, A. J. (1963). Transient voice changes associated with emotional stimuli. *Archives of General Psychiatry*, *8*, 362–365.

Bachorowski, J., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, *6*, 219–224.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality & Social Psychology*, *70*, 614–636.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*, 276–298.

Cacioppo, J. T., Klein, D. J., Bernston, G. G., & Hatfield, E. (1993). The psychophysiology of emotion. In M. Lewis & J. Haviland (Eds.), *Handbook of emotion* (pp. 119–142). New York: Guilford Press.

Childers, D. G., & Krishnamurthy, A. K. (1985). A critical review of electroglottography. *CRC Critical Reviews in Biomedical Engineering*, *12*, 131–161.

Dawson, M. E., Schell, A. M., & Filion, D. L. (1990). The electrodermal system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 295–324). New York: Cambridge University Press.

Deller, J. R., Proakis, J. G., & Hansen, J. H. L. (1993). *Discrete time processing of speech signals*. New York: Macmillan.

Duncan, G., Laver, J., & Jack, M. A. (1983). *A psycho-acoustic interpretation of variations in divers' voice fundamental frequency in a pressured helium-oxygen environment*. Work in Progress Report Ser. No. 16 (S.9–16), University of Edinburgh, Department of Linguistics.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fant, G. (1993). Some problems in voice source analysis. *Speech Communication*, *13*, 7–22.

Fourcin, A. J., & Abberton, E. (1971). First applications of a new laryngograph. *Medical Biology Illustustrated*, *21*, 172–182.

Fourcin, A. J., & Abberton, E. (1977). Laryngograph studies of vocal-fold vibration. *Phonetica*, *34*, 313–315.

Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Henrich, N., d'Alessandro, C., Doval, B., & Castellengo, M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *Journal of the Acoustical Society of America*, *115*, 1321–1332.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*, 887–900.

Johnstone, T., van Reekum, C. M., Hird, K., Kirsner, K., & Scherer, K. R. (2005). Affective speech elicited with a computer game. *Emotion*, *5*, 513–518.

Johnstone, T., Van Reekum, C. M., & Scherer, K. R. (2001). Vocal expression correlates of appraisal processes. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 271–284). Oxford: Oxford University Press.

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, *1*, 381–412.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Kanouse, D. E., & Hanson, L. R. J. (1987). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, & S. Valins (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 47–62). Hillsdale, NJ: Erlbaum.

Kappas, A., & Pecchinenda, A. (1999). Don't wait for the monsters to get you: A video game task to manipulate appraisals in real time. *Cognition & Emotion*, *13*, 119–124.

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *Forensic Linguistics*, *4*, 104–124.

Laukka, P., Juslin, P. N., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, *19*, 633–653.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

MacDowell, K. A., & Mandler, G. (1989). Constructions of emotion: Discrepancy, arousal, and mood. *Motivation and Emotion*, *13*, 105–124.

Marasek, K., & Pützer, M. (1997). *Electroglottographical differentiation of pathological voice qualities*. *LARYNX-1997*. Marseille, France. pp. 91–94.

Nikula, R. (1991). Psychological correlates of nonspecific skin conductance responses. *Psychophysiology*, *28*, 86–90.

Öhman, A., Esteves, F., Flykt, A., & Soares, J. J. F. (1993). Gateways to consciousness: Emotion, attention, and electrodermal activity. In J. -C. Roy, W. Boucsein, D. C. Fowles, & J. H. Grutelier (Eds.), *Progress in electrodermal research* (pp. 137–157). New York: Plenum.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*, 143–165.

Scherer, K. R., Johnstone, T., & Klasmeyer, G. (2003). Vocal expression of emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433–456). New York: Oxford University Press.

Simonov, P. V., & Frolov, M. V. (1973). Utilization of human voice for estimation of man's emotional stress and state of attention. *Aerospace Medicine*, *44*, 256–258.

Smith, G. A. (1977). Voice analysis for the measurement of anxiety. *British Journal of Medical Psychology*, *50*, 367–373.

Sundberg, J. (1995). Vocal fold vibration patterns and modes of phonation. *Folia Phoniatrica Et Logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *47*, 218–228.

Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *The Journal of the Acoustical Society of America*, *96*, 2101–2107.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *65*, 248–260.

van Reekum, C. M., Johnstone, T., Banse, R., Etter, A., Wehrle, T., & Scherer, K. R. (2004). Psychophysiological responses to appraisal dimensions in a computer game. *Cognition & Emotion*, *18*, 663–688.

(Received January 5, 2006; Accepted April 29, 2007)