



Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep



Understanding emotional expression using prosodic analysis of natural speech: Refining the methodology

Alex S. Cohen^{a,*}, S. Lee Hong^b, Alvaro Guevara^c

^a Louisiana State University, Department of Psychology, 236 Audubon Hall, Baton Rouge, LA 70803, USA

^b Indiana University, Department of Kinesiology, USA

^c Louisiana State University, Department of Mathematics, USA

ARTICLE INFO

Article history:

Received 21 April 2009

Received in revised form

27 November 2009

Accepted 27 November 2009

Keywords:

Emotion
Prosody
Affect
Expression
Speech
Mood
Laboratory
Schizophrenia
Depression

ABSTRACT

Emotional expression is an essential function for daily life that can be severely affected in some psychological disorders. Laboratory-based procedures designed to measure prosodic expression from natural speech have shown early promise for measuring individual differences in emotional expression but have yet to produce robust within-group prosodic changes across various evocative conditions. This report presents data from three separate studies (total $N = 464$) that digitally recorded subjects as they verbalized their reactions to various stimuli. Format and stimuli were modified to maximize prosodic expression. Our results suggest that use of evocative slides organized according to either a dimensional (e.g., high and low arousal – pleasant, unpleasant and neutral valence) or categorical (e.g., fear, surprise, happiness) models produced robust changes in subjective state but only negligible change in prosodic expression. Alternatively, speech from the recall of autobiographical memories resulted in meaningful changes in both subjective state and prosodic expression. Implications for the study of psychological disorders are discussed.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Emotional expression is essential to a wide range of human functions (Decety & Lamm, 2006; Gross, 2002; LeDoux, 2000) and is compromised in a host of psychiatric and developmental disorders, such as schizophrenia (Cohen, Alpert, Nienow, Dinzeo, & Docherty, 2008; Cohen et al., 2005), depression (Leventhal, Chason, Tapia, Miller, & Pettit, 2006) and autism (Matese, Matson, & Sevin, 1994; South et al., 2008). Objectively measuring emotional expression has been an integral endeavor for a wide range of empirical pursuits. One particularly important method involves acoustic analysis of vocal expression. Despite considerable empirical work employing acoustic analysis of natural speech to understand expressivity in both pathological and nonpathological samples (Scherer, 2003; Sobin & Alpert, 1999), the field has been limited by the lack of a standardized method of speech procurement.

The Computerized Assessment of affect from Natural Speech (CANS; Cohen, Minor, Najolia, & Hong, 2009) is a laboratory-based procedure designed to measure both lexical and prosodic expression from natural speech across a range of evocative conditions. The CANS procedure involves having individuals generate free-speech in response to standardized stimuli. In this manner, the procedure is highly controllable in terms of mood-induction effects and is repeatable, while also being sensitive and applicable to a wide range of both clinical and non-clinical populations. This methodology also allows for generation of free-speech for lexical examination (see Cohen, Iglesias, & Minor, 2009; Cohen, Minor, et al., 2009 for expansion on this point). Several recent studies have offered preliminary support for the CANS. First, two studies of adults found that prosodic expression generated in response to affectively valenced pictures from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005) significantly changed across neutral, pleasant and unpleasant conditions. These CANS variables showed high temporal stability, suggesting they are a reliable index of individual differences in emotional expression. Changes in CANS variables also corresponded to subjective appraisals of state arousal but not valence (i.e., pleasant or unpleasant mood ratings). This suggests that prosodic expression

* Corresponding author. Tel.: +1 225 578 7017.

E-mail address: acohen@lsu.edu (A.S. Cohen).

reflects relatively specific nervous system activities that are distinct from subjective emotional states.

Although the CANS methodology appears promising for a wide range of applications, the procedure is limited in that the magnitude of prosodic change, while statistically significant, has been marginal. Note that in the aforementioned CANS studies, the effect sizes between conditions were generally in the negligible range (i.e., <0.20 ; Cohen, 1988). Thus, one can question the degree to which the CANS can elicit meaningful prosodic changes. This may be a reflection, in part, of the fact that it was subjective arousal, not valence ratings that were associated to prosodic change, even though the stimuli were categorized solely on valence (neutral, pleasant, and unpleasant). This becomes an issue considering that valence and arousal are not necessarily linked (Cacioppo & Berntson, 1999). That is, visual stimuli with pleasant and unpleasant tones can have dramatically different arousal ratings. For example, a “cemetery scene” versus “mutilated bodies” would yield low and high arousal levels respectively, although both stimuli would be categorized as unpleasant.

We thus turn to the circumplex model of emotion as a model for understanding variability in emotions – wherein emotional states are viewed as occupying a two-dimensional valence \times arousal space. This is a well-established model validated for psychological research that may be more appropriate for conceptualizing prosodic change in that it accounts for both valence and arousal simultaneously (Russell, 1980, 2009). In the first study presented here, our goal was to maximize prosodic expression across conditions. We thus administered a CANS procedure to healthy adults using stimuli that evoke separate high and low arousal with neutral, pleasant and unpleasant valences.

While the circumplex model has been one of the primary concepts in fundamental emotion research to date, there is surprisingly little research on valence and arousal in terms of prosody. Rather, most prior prosody studies have conceptualized emotions using categorical models comprising distinct emotions. While a variety of models exist (e.g., Frijda, Kuipers, & ter Schure, 1989; Scherer, 1984; “natural kinds” model, Barrett, 2006; Ekman, Friesen, & Ancoli, 1980) and there are differences in putative neurobiological or cognitive underpinnings across these models, a commonality across them is that emotions reflect discrete categories. Using this categorical taxonomy, different emotional states are reflected as categorically distinct “kinds” – anger, fear, surprise, happiness. Prosodic changes across these emotional states have been well documented (see Sobin & Alpert, 1999 for a review). For example, anger and fear have been associated with increases in inflection, amplitude and emphasis while sadness and disgust have been associated with declines in at least some of these variables (Scherer, 2003; Sobin & Alpert, 1999; Vervididis & Kotropoulos, 2006). In a second study presented here, we employed a CANS procedure using stimuli that represent a range of distinct emotional states.

With respect to maximizing prosodic changes across conditions, a final point to consider is the effectiveness of picture stimuli in evoking prosodic changes. While picture stimuli can effectively evoke a range of subjectively-reported emotional states (Lang et al., 2005), these ratings are not immune to demand characteristics. Thus, the degree to which emotion is actually experienced is unclear, especially since it is not strongly observed in prosodic change. A particular concern is that picture stimuli may not provide sufficient personal relevance for an individual. An alternative to using picture stimuli is to have subjects offer freely-generated autobiographical statements. While not as objective as picture stimuli, a free-recall condition may elicit stronger prosodic reactions because the task is more personally relevant. The third study presented here employs autobiographical free-recall to evoke prosodic changes.

In summary, the present study indirectly compared three separate CANS formats to determine which was optimal for maximizing prosodic expression across conditions. We thus employed three experiments with the following conditions: (1) picture stimuli based on a dimensional “Circumplex” model of emotion (Russell, 1980); (2) picture stimuli based on a categorical model (Barrett, 2006; Ekman et al., 1980); and (3) free-recall of autobiographical memories. Our method of evaluating these procedures is based on comparing prosody variables across the various conditions based on both statistical significance and magnitude of effect, using effect size statistics.

2. Study 1

2.1. Materials and methods

2.1.1. Participants

Subjects were 56 males and 155 females recruited from a large public university in the Southeastern United States. The average age was 20.75 years (standard deviation (SD) = 3.74), and all subjects were native English speakers and reported having 20/20 (or corrected to) vision. Subjects received course credit for participating in this experiment. This study was approved by the appropriate Institutional Review Board and all subjects provided written informed consent prior to beginning the study.

2.1.2. Procedures

Subjects were seated in front of a computer monitor out of view of laboratory assistants. The experiment was run using Eprime software version 2.0 (Psychology Software Tools, 2002). Subjects were asked to view five separate randomly-ordered blocks of five affectively Unpleasant-high arousal (picture numbers 5971, 1300, 6550, 3010, 6230), Unpleasant-low arousal (picture numbers 9220, 9265, 2722, 9001, 9331), Pleasant-high arousal (picture numbers 8185, 8030, 8490, 8370, 8502), Pleasant-low arousal (picture numbers 7175, 7010, 7004, 7490, 7187), Neutral-high arousal (picture numbers 8160, 8179, 5950, 1931, 1321) and Neutral-low arousal (picture numbers 5000, 2370, 5010, 1610, 5200) pictures from the International Affective Picture System (Lang et al., 2005).

Picture display was 20 s such that the total time allotted for speech was 100 s. Blocks were separated by a 10 s interval during which subjects were asked to “relax and breathe deeply”. We assumed that this epoch was sufficient to facilitate a return to baseline emotion levels based on prior studies reporting that individuals return to electrophysiological baseline after processing IAPS stimuli within 1 s of stimulus offset (Lang, Bradley, & Cuthbert, 1999). After each picture block, subjects separately rated their arousal level, pleasant emotion and unpleasant emotional state using a modified Self-Assessment Manikin scale (SAM; Lang et al., 2005) ranging from one (neutral emotion and high arousal respectively) to nine (pleasant, unpleasant emotion and low arousal respectively). SAM ratings are based on the Circumplex model of emotion (Russell, 1980), although we separated pleasant and unpleasant ratings based on evidence that these dimensions are not necessarily linked (Cacioppo & Berntson, 1999).

During the voiced condition, subjects were asked to verbalize their thoughts about the picture, notably how it made them feel and what memories it conjured. Their speech was digitally recorded at 16 bits per second at a sampling frequency of 44.1 kHz using a headset microphone. Prior to each block of pictures, the experimenter read the following instructions.

“In a moment, I will show you a series of pictures for a minute and a half. While you are focusing on these pictures, I want to record you as you talk to me about these pictures. I want to know how

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات