

EXAMINING VOCAL AFFECT IN NATURAL VERSUS ACTED EXPRESSIONS OF EMOTION

Barbra Zupan¹ and Madeline Shiah²

¹Department of Applied Linguistics, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, Canada L2S 3A1

²Department of Communication Disorders and Sciences, University at Buffalo, 122 Cary Hall, Buffalo, NY 14214

1. INTRODUCTION

Vocal Affect (VA) is one of the most significant characteristics by which human emotion is identified. This non-linguistic component of speech includes prosody, intonation, speaking rate, and vocal effort; all of which aid in conveying emotional information beyond the semantic content of a message.¹ Although listeners pay attention to both non-linguistic and linguistic properties of the speaker's utterance to derive meaning from the message, a large amount of information can be drawn from vocal affect information alone.^{1,2} The average listener is generally proficient in detecting a communicative partner's emotional state based on the associated acoustic properties of a happy, sad, or angry tone of voice. These acoustic properties are important since the intention of a speaker's message may change significantly depending on which emotional tone of voice is used. Thus, a listener must combine what is said with how it is spoken to accurately gauge the meaning of the speaker's message.

During a typical face-to-face conversational exchange, a listener is able to use numerous sources of information to discern the intended meaning of the speaker including facial, gestural and vocal cues. In fact, successful social interactions rely on the integration and interpretation of facial and vocal affect cues.² However, some real-life situations, such as having a telephone conversation or in the case of visual deficits, force one to rely on only VA. This type of scenario was of particular interest to us. Thus, the aim of the current investigation was to examine a person's ability to recognize emotion in a vocal message in the absence of facial and gestural cues. It was hypothesized that listeners would be able to identify natural expressions of emotion more accurately than acted portrayals of emotion due to the physiological changes said to underlie the vocal qualities of the speakers in natural emotional expressions. In addition, we would assume that natural expressions of emotion are more typically heard, experienced, or encountered, therefore should be more easily identifiable.

2. METHOD

2.1 Participants

Phase One included seven Brock University students as speakers (4 female, 3 male), ranging in age from 21-25.

Phase Two included 31 Brock University students as listeners (27 female, 4 male), with a mean age of 21.8 years. All participants spoke English as their primary language,

had no hearing or uncorrected visual deficits, and reported no current or previous communication disorders.

2.2 Procedure

Phase One: The examiners and a research assistant viewed DVDs of reality television shows ("Punk'd", "Amazing Race", & "The Real World") for the purpose of selecting clips of characters speaking semantically neutral phrases using the emotions fear, anger, and happiness (5 clips of each emotion). Only clips identified by the examiners as being significantly characteristic of the tone of voice associated with the selected emotions were selected. To create the acted stimuli, speakers were asked to say the same phrases as those selected from the reality shows, in the same emotional tone(s) of voice. Speakers were recorded using an Olympus digital audio recorder with consistent mouth to microphone distance.

Phase Two: Listeners were administered the vocal affect portion of the Diagnostic Analysis of Nonverbal Affect-2 (DANVA-2). This test requires participants to listen to 24 trials of a sentence and identify the speaker's tone of voice (Happy, Sad, Angry, Fearful). The vocal affect portion of the DANVA-2 served as a baseline for participants' general ability to identify emotional tones of voice. Following completion of the DANVA-2, participants were seated comfortably in front of a computer and instructed to listen to the test stimuli (natural and acted audio clips) and identify as accurately as possible which of the three emotions was being portrayed in each of the vocal clips. Participants were also asked to indicate whether they thought the expression was a natural or acted (pretended) portrayal. Stimuli were randomized for each participant using Cedrus SuperLab software. Responses were provided using a 6 button response pad with the following options: Happy; Happy Acted; Angry; Angry Acted; Fearful; and Fearful Acted. Listeners received no feedback regarding whether the stimuli were from the natural or acted conditions or if the selected emotion was correct.

3. RESULTS

A repeated measures ANOVA revealed a significant interaction between acted versus natural emotion showing better identification of acted than natural expressions of emotion [$F(2,29)=18.277$ $p<.01$]. Follow-up paired samples t test were conducted with alpha set at .01 to control for Type 1 error, to evaluate participants' ability to correctly identify natural and acted expressions within emotion type (see Figure 1). The results showed that both Happy Acted

and Fearful Acted stimuli were identified with significantly higher accuracy better than their Natural counterparts. Results of the identification of Angry however were the opposite, with natural expressions of Angry being significantly better identified by participants than acted stimuli.

Overall, participants were better able to identify acted versus natural expressions of emotion. Since the DANVA-2 includes acted vocal expressions, we computed a correlation analysis between participants' DANVA-2 score and their performance in identifying acted expressions of the three emotion types. No significant correlation was found, $r(29) = .32, p = <.085$.

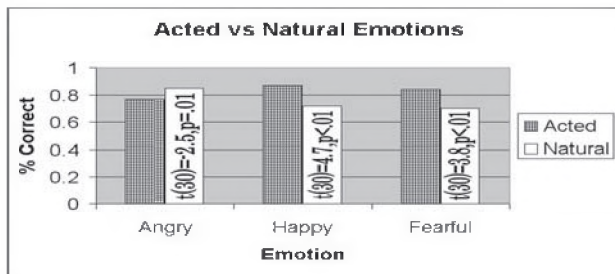


Figure 1. Accuracy of identification of acted vs. natural expressions.

4. DISCUSSION

The current study aimed to examine the ability of listeners to identify vocal expressions of emotion (Happy, Angry, Fearful) under vocal-only conditions in both natural and acted conditions. It was hypothesized participants would more accurately identify natural expressions of emotions since these expressions have been associated with underlying physiological changes that affect the acoustic pattern of the voice.³ This assumption is an important one since equal accuracy in both categories (natural and acted) would indicate that these physiological changes are not a necessity in the identification of vocal expressions. Our hypothesis was confirmed for Angry vocal expressions, but not for Happy and Fearful vocal expressions. Interestingly, the vocal expressions portrayed in the acted portrayals were more exaggerated expressions of these two emotions while the natural expressions of Anger contained more prominent acoustic cues. Thus, it appears that participants may perceive more prominent acoustic cues as acted because they are more exaggerated than the emotion expressions they encounter in typical daily interactions. Participants in this study were aware that there were examples of both acted and natural expressions and therefore, may have used exaggerated versus subtle expressions in their decision making process. This was reflected in the type of response participants selected since they tended to label more obvious expressions as acted on the response pad. However, these results need to be interpreted cautiously since the number of acted stimuli in comparison to natural stimuli in the current study was unbalanced (105 acted clips; 15 natural clips). The minimal number of natural clips reflects the challenge of finding semantically neutral phrases that portray one of

the three emotions used in this study within reality television shows. The use of unequal acted versus natural expressions may have masked inaccuracies of participants simply due to the fact that they had a better chance overall at correctly identifying an acted clip since there were more of them. In addition, participants may have been expecting a more equal balance which may have influenced their choice. This was certainly a limitation of the current study.

A lack of correlation between responses to the DANVA-2 and acted vocal expressions was surprising, since both sets of stimuli focus on acted portrayals. However, the current study included only portrayals of angry, fearful, and happiness whereas the DANVA-2 also includes portrayals of sadness. Due to the increased jitter and breathiness in both fearful and sadness, listeners often confuse these two emotions. Since the current study did not include sadness, the lack of correlation may simply reflect this difference.

Future work in this area should also examine the acoustic cues of the acted versus natural emotion expressions. Although the use of natural clips from reality television shows provided us with stimuli more likely to represent the acoustic patterns associated with physiological changes for each emotion category, we were unable to analyze the stimuli due to the presence of background noise and music present in the majority of clips. Since we could not remove the music, we resolved the issue for the current study by adding similar background music to the acted stimuli to ensure that responses were not being influenced by this factor. However, spectrogram analysis of these two types of stimuli remains important to examine whether the acoustic cues associated with each emotion are influenced by this condition.

REFERENCES

1. Nygaard, L.C. & Queen, J.S. (2008). Communicating emotion: linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 1017-1039.
2. Zupan, B., Neumann, D., Babbage, D.R., & Willer, B. (2009). The importance of vocal affect to bimodal processing of emotion: implications for individuals with traumatic brain injury. *Journal of Communication Disorders*, 42(1), 1-17.
3. Johnstone, T., & Scherer, K.R. (2003). Vocal communication of emotion. In M.L. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed., pp 220-235). New York: The Guilford Press.