Brain and Cognition, **48**(2-3), 499-514 (2002)

Evaluation of nonverbal emotion in face and voice: some preliminary findings on a new battery of tests

Marc David Pell
McGill University, Montréal

Abstract

This report describes some preliminary attributes of stimuli developed for future evaluation of nonverbal emotion in neurological populations with acquired communication impairments. Facial and vocal exemplars of six target emotions were elicited from four male and four female encoders and then pre-judged by 10 young decoders to establish the category membership of each item at an acceptable consensus level. Representative stimuli were then presented to 16 additional decoders to gather indices of how category membership and encoder gender influenced recognition accuracy of emotional meanings in each nonverbal channel. Initial findings pointed to greater facility in recognizing target emotions from facial than vocal stimuli overall, and revealed significant accuracy differences among the six emotions in both the vocal and facial channels. The gender of the encoder portraying emotional expressions was also a significant factor in how well decoders recognized specific emotions (*disgust, neutral*), but only in the facial condition.

Introduction

Nonverbal displays provide a rich source of cues regarding an individual's affective intentions in speech or during other interpersonal events. Although nonverbal symbol systems serve a variety of social-pragmatic goals of the 'actor' (*encoder*) by influencing how 'partners' (*decoders*) interpret the communicative situation, the cognitive operations supporting the facial or vocal processing of emotion are frequently studied in reference to pan-culturally recognizable, 'primary' emotional states (e.g., Ekman, 1992; Ekman & Friesen, 1976). Thus, deficits in the ability to perceive, label, and/or otherwise gauge the significance of facial or vocal-prosodic stimuli representing *happiness*, *sadness*, *anger*, or other basic emotions have been uncovered following various forms of cerebral pathology (Blonder, Bowers, & Heilman, 1991; Borod, Cicero, & Obler et al., 1998; Pell, 1996; 1998). It is largely assumed that such impairments are associated with a decline in communicative competence and social prowess which may be amenable to intervention; however, first, detailed diagnostic tools capable of illuminating the types of emotional impairments present in individuals with different forms of neurological compromise must be developed and refined.

The present study examines important perceptual attributes of a subset of stimuli constructed for future evaluation of nonverbal emotional communication in brain-damaged populations. Specifically, facial and vocal stimuli representing six discrete emotional categories, including positive, negative, and "neutral" exemplars in each nonverbal channel, were elicited from both male and female encoders for perceptual evaluation by healthy control subjects. Based on past research in this area, it was expected that differences in the meaning among specific emotions entered into the

dataset, and the gender of the encoder posing the emotions, were important variables to monitor in characterizing the perceptual dimensions of these stimuli (Ekman & Friesen, 1976; Scherer, Banse, Wallbott, & Goldbeck, 1991; Sobin & Alpert, 1999). The perceptual attributes of related nonverbal stimuli conveying *non-affective* meanings elicited from the same group of encoders were also of interest but are not reported here (Pell, in preparation).

Methods

1) Stimulus construction and validation study

Materials - Stimuli consisted of static poses of facial expressions (facial condition) or short English utterances in which key words were substituted with speech-like but meaningless elements that rendered the sentence nonsensical but an appropriate carrier of suprasegmental information (vocal condition; see Scherer et al., 1991 for related stimuli). In each nonverbal channel, stimuli were encoded by several individuals to represent a range of 'primary' emotional meanings (Ekman, 1992), balanced roughly for the positive or negative *valence* of the expression (Borod et al., 1998). Each series included two positive emotions (*happiness, pleasant surprise*), three negative emotions (*anger, sadness, disgust*), and *neutral* displays, for which actors attempted to background affective features in the visual or auditory signal (see Young, Rowland, & Calder et al., 1997 for a discussion of the "neutral" category).

Subjects/encoding procedure - Four male and four female 'encoders' with a background in amateur theatre were recruited to portray both facial and vocal depictions of target emotions. Encoders varied in age (female: 20-58 yrs.; male: 22-62 yrs.) and were recorded during a single 2-3 hour session, beginning with the facial condition and terminating with the vocal condition. In the facial condition, encoders posed each of the six emotional expressions in sequence accompanied by examiner feedback to guide their performance (e.g., description of circumstances that may elicit the target emotion, of associated muscle movements)(Borod et al., 1998; Ekman & Friesen, 1976). Repeated poses of each target were recorded continuously by a digital camcorder. Video recordings were later captured and edited in Adobe Premiere to extract at least ten still exemplars of each target expression as colour bitmaps (480 x 640). In the vocal condition, actors were prompted to say and repeat five nonsense utterances in each target emotion (e.g., *Someone migged the pazing*), again accompanied by cues from the examiner. Vocal recordings were captured onto digital audiotape through a professional headmounted microphone and then transferred and edited on a PC to select the best two representations of each emotion, per item and actor. Collectively, this exercise yielded over 500 facial and vocal stimuli conducive to perceptual evaluation in a validation study.

Emotion validation procedure - It was anticipated that many of the emotional stimuli recorded would not consistently portray one of the desired targets when evaluated by decoders, or that other nuances of emotional meaning would be identifiable with particular items (Ekman & Friesen, 1976; Scherer et al., 1991). A validation procedure was therefore essential to eliminate stimuli in which encoders' portrayal did not adequately communicate one of the intended emotional targets to decoders. All items were presented to ten healthy English speakers (5 female, 5 male) in two separate labelling tasks: facial and vocal. In each condition, stimuli were presented individually in blocks of approximately 30 via laptop computer. Each stimulus was judged by the

decoder as belonging to one of the six target emotions (*happy, surprised, angry, sad, disgusted, neutral*) or as conveying a different emotional quality, in which case participants were permitted to write their response on a checklist. Based on the group accuracy for each item, and owing to the "open" response paradigm of the validation study, a liberal consensus rate of 60% correct target responses was adopted to limit further analyses in each condition to only those stimuli highly representative of a single intended emotion. As reported in the Table, this exercise yielded 129 facial stimuli (X=21.5/emotion) and 91 vocal stimuli (X=15.2/emotion) suitable for detailed specification of perceptual features in the emotion recognition study.

2) Emotion recognition study

Subjects - Pre-validated facial and vocal stimuli were presented to a new group of young healthy decoders (undergraduate students at McGill). To date, data on 10 female and 6 male subjects contribute to the reported findings.

Evaluation procedure & analyses – The six target emotions were presented for identification by decoders in a manner largely similar to that described in the validation procedure. Decoders were allowed 10 seconds to judge the emotion of a still photograph (facial condition) or they listened to emotionally-inflected nonsense sentences presented over headphones (vocal condition). Recognition of each stimulus were indicated by clicking on the verbal label, listed in a fixed random order at the bottom of the computer screen, that corresponded to its emotional meaning from the six target emotions. Experimental subjects did not have recourse to alternative emotion designations as was permitted in the validation procedure. Response accuracy was then examined to explore the concurrent influence of emotion category membership and encoder gender on emotion judgements in each separate nonverbal channel using 6 x 2 repeated measures ANOVAs (posthoc comparisons were accomplished using Tukey's HSD procedure). Preliminary data do not yet permit commentary on how gender of the d*ecoder* impacts on perceptual performance in reference to the chosen variables (Scherer et al., 1991).

Results

Emotion recognition as a function of category membership

Stimulus-response matrices depicting how the 16 decoders responded to exemplars of each emotional category in the facial and vocal conditions are supplied by the Table. It may be observed that decoders were less accurate in identifying emotional characteristics of vocal than facial stimuli posed by the same group of encoders, a trend that was true for each target emotion investigated here. The Table further reveals that in each nonverbal channel, displays of *happiness* were associated with the least amount of confusion with potential foils, whereas *neutral* displays were recognized most poorly in both conditions and were frequently attributed various emotional qualities by decoders. *Pleasant surprise* was identified consistently in the facial channel but far less reliably when communicated vocally.

Results of separate 2 x 6 ANOVAs confirmed that there were significant differences in how well decoders recognized the six target emotions in each nonverbal channel [EMOTION main effect: $F_{FACIAL}(5,75) = 9.07$, $p<.001$; $F_{VOCAL}(5,75) = 2.46$, $p<.05$]. In the facial condition, differences in emotion recognition were further dependent on the gender of the actor [EMOTION x GENDER: $F_{FACIAL}(5,75) = 6.14$, $p<.001$] and are

discussed below. Evaluation of *vocal* emotion was not dependent on gender specifications of the encoder, and in fact, was associated with relatively few differences among the six emotions altogether: *happiness* was recognized significantly better than *pleasant surprise* based on judgements of vocal-prosodic cues. No other comparisons among the six emotion means proved significant ($p>.05$).

[INSERT TABLE ABOUT HERE]

Emotion recognition as a function of actor gender

Although gender of the encoder portraying emotional displays did not independently affect recognition performance in either communicative channel, this factor did influence how decoders judged specific emotional qualities in the facial channel only. Correct attributions of *neutral* affect were significantly fewer when posed by female than male encoders; for *disgust*, the opposite pattern was true, as decoders recognized these qualities in female faces more accurately than in male faces. Viewed another way, when the encoder was female, decoders correctly labelled *neutral* significantly less frequently than all other emotional designations (which did not differ). When the encoder was male, however, *neutral* was recognized significantly less than *happiness* only (happy faces were associated with a ceiling effect for both male and female encoders). The relationship between emotion and encoder gender in the facial condition is illustrated by the Figure.

[INSERT FIGURE ABOUT HERE]

Discussion

The ability to infer emotion from vocal and facial signal systems is basic to human communication and interaction. The present serves as an initial report on how young healthy decoders judge emotion as a function of nonverbal channel and gender of the encoder for a newly-developed corpus of stimuli. As predicted (Borod, Pick, Hall et al., 2000; Wallbott & Scherer, 1986), recognition of primary emotions was less accurate overall for vocal-prosodic stimuli than for facial stimuli portrayed by the same encoders (X=78% versus 91% correct in the vocal and facial conditions, respectively). Also, it merits emphasizing that a substantially greater number of "valid" exemplars of each target emotion, which did not convey non-target emotions or blends of different emotional qualities, emerged in the facial condition than in the vocal condition following the validation procedure (with the notable exception of *sad*-review Table).

In the vocal channel, early findings indicate that reliability of emotion judgements is not systematically influenced by the gender of the encoder (Wallbott & Scherer, 1986). This conclusion is warranted despite clear evidence that *individual* encoders of both genders differed considerably in their ability to portray individual emotions successfully to decoders (see Pell, in preparation; also Scherer et al., 1991). Recognition of prosodic indicators of *pleasant surprise* and *happiness* was noteworthy in this condition; *pleasant surprise* was recognized most poorly overall (65% correct), albeit well above chance (16.7%). In the case of *happiness*, this emotion was associated with a very small number of highly recognizable exemplars (i.e., the fewest *happy* stimuli met criterion in the validation study, but these items were then most accurately identified by the 16 decoders in the vocal condition of the recognition study). Relative difficulties decoding both *happiness* and *surprise* vocally has been reported previously (Wallbott & Scherer, 1986)

and may highlight more extensive overlap between these two categories in the present dataset. Finally, it is interesting that the present decoders experienced little difficulty recognizing the vocal correlates of *disgust*, contrary to some prior findings in this literature (Scherer et al., 1991).

Recognition performance in the facial condition, which was highly accurate overall, was nonetheless contingent on the emotion being decoded and the gender of the individual portraying the emotion. Specifically, *disgust* was judged more accurately when posed by female than male encoders. Moreover, *neutral*—which is thought to represent a discrete category distinct from 'basic' target emotions investigated here, with characteristic facial markers (e.g., Young et al., 1997)—was identified significantly better when posed by male than female encoders. It is not immediately clear whether the influence of encoder gender on these two emotions reflects systematic properties of how these emotions are decoded and labelled, or whether certain artefacts may have been introduced by the small number of encoders contributing to this database to date (comparative data on *neutral*, in particular, is lacking). The interaction between encoder gender and these two emotions is thus interpreted cautiously, but merits closer attention through studies that employ a larger number of encoders and decoders than participated in the current offering.

There is little doubt that acquired disturbances in the recognition of either facial or vocal nonverbal information can yield profound deficits in social-pragmatic skills and interpersonal competence following brain insult. The construction of tests suitable for detailed evaluation of nonverbal processes in neurologically-impaired adults, which tap both emotional and non-emotional meanings of nonverbal stimuli (e.g., Blonder et al., 1991; Borod et al., 1998), thus constitutes an important step in the identification and eventual remediation of nonverbal communicative impairments in affected brain-damaged individuals. Continued development and validation of the *McGill Assessment of Emotional Communication* will help achieve these ends.

References

Blonder, L., Bowers, D., & Heilman, K. (1991). The role of the right hemisphere in emotional communication. Brain, *114* (3), 1115-1127.

Borod, J., Cicero, B., Obler, L., Welkowitz, J., Erhan, H., Santschi, C., Grunwald, I., Agosti, R. & Whalen, J. (1998). Right hemisphere emotional perception: evidence across multiple channels. Neuropsychology, *12* (3), 446-458.

Borod, J., Pick, L., Hall, S., Sliwinski, M., Madigan, N., Obler, L., Welkowitz, J. Canino, E., Erhan, H., Goral, M., Morrison, C. & Tabert, M. (2000). Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. Cognition and Emotion, *14* (2), 193-211.

Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, *6*, (3/4), 169-200.

Ekman, P. & Friesen, W. (1976). Pictures of facial affect. Palo Alto, CA: Consulting Psychologists Press.

Pell, M. (in preparation). Judgement of emotion in nonverbal signals.

Pell, M. (1996). On the receptive prosodic loss in Parkinson's disease. Cortex, *32* (4), 693-704.

Pell, M. (1998). Recognition of prosody following unilateral brain lesion: influence of functional and structural attributes of prosodic contours. Neuropsychologia, *36* (8), 701-715.

Scherer, K., Banse, R., Wallbott H., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. Motivation and Emotion, *15* (2), 123-148.

Sobin, C. & Alpert, M. (1999). Emotion in speech: the acoustic attributes of fear, anger, sadness, and joy. Journal of Psycholinguistic Research, *28* (4), 347-365.

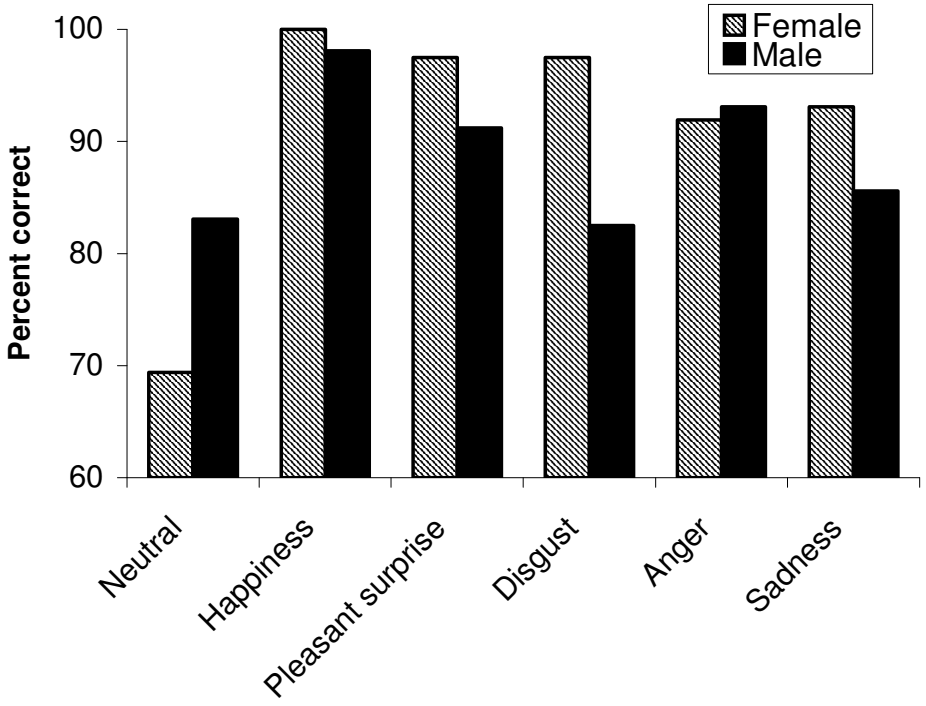
Wallbott, H. & Scherer, K. (1986). Cues and channels in emotion recognition. Journal of Personality and Social Psychology, *51* (4), 690-699.

Young, A., Rowland, D., Calder, A., Etcoff, N., Seth, A., & Perrett, D. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition.

TABLE. Recognition of six emotions posed in the facial and vocal channel by eight encoders when judged by 16 healthy decoders (percentage of total responses).

		RESPONSE					
TARGET EMOTION	***n***	Neutral	Happiness	Pleasant surprise	Disgust	Anger	Sadness
Facial Stimuli =	129						
Neutral	21	**80**	4	0	1	5	10
Happiness	24	1	**99**	0	0	0	0
Pleasant surprise	21	0	4	**95**	0	0	1
Disgust	22	0	2	0	**91**	1	6
Anger	22	1	1	0	3	**92**	2
Sadness	19	4	0	0	7	0	**88**
Vocal Stimuli =	91						
Neutral	11	**73**	9	1	0	1	16
Happiness	9	5	**86**	6	1	1	1
Pleasant surprise	16	2	27	**65**	4	3	0
Disgust	13	5	1	7	**77**	6	4
Anger	13	3	0	1	12	**83**	0
Sadness	29	15	0	0	2	1	**82**

FIGURE. Recognition of each target emotion portrayed in the facial channel as a function of the gender of the encoder.



Contact information:
Marc Pell, Ph.D.
McGill University
School of Communication Sciences & Disorders
1266 ouest, avenue des Pins
Montréal (Qc) H3G 1A8
tel: (514) 398-4133
email: *marc.pell@mcgill.ca*