

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

EMOTIONAL PROSODY PRODUCTION IN BRAIN-DAMAGED POPULATIONS:
HUMAN JUDGMENTS AND ACOUSTICAL ANALYSIS

by

J. MICHAEL SCHMIDT

A dissertation submitted to the Graduate Faculty in Psychology (Neuropsychology
Subprogram) in partial fulfillment of the requirements for the degree of Doctor of
Philosophy, The City University of New York

2003

UMI Number: 3103167

UMI[®]

UMI Microform 3103167

Copyright 2003 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Psychology (Neuropsychology Subprogram) in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

September 3, 2003
Date

Joan P. Borod
Chair of Examining Committee

September 4, 2003
Date

[Signature]
Executive Officer

Joan Borod, Ph.D.

Howard Ehrlichman, Ph.D.

Nancy Foldi, Ph.D.
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract**EMOTIONAL PROSODY PRODUCTION IN BRAIN-DAMAGED POPULATIONS:
HUMAN JUDGMENTS AND ACOUSTICAL ANALYSIS**

by

J. Michael Schmidt

Adviser: Joan C. Borod, Ph.D.

This study investigated two areas of emotion: acoustical parameters used to communicate discrete emotions and brain structures involved in the expression of prosodic emotion. Further, the right-hemisphere, valence, motoric direction, and cue-lateralization emotional communication hypotheses were tested. Speech samples (n = 796) produced by 20 right brain-damaged, 17 left brain-damaged, and 18 demographically matched normal control adult men and women were selected from happy, sad, angry, and neutral monologues, low-pass filtered to eliminate language cues, and presented to 14 naïve judges who rated the speech samples for emotional intensity and emotion-type identification. Additionally, these samples underwent computerized acoustical analysis to measure pitch, duration, and volume. Analyses revealed that pitch mean and standard deviation, volume standard deviation, and speech rate were greatest for the expression of happiness and smallest for sadness. Anger was also associated with high pitch mean and standard deviation level, as well as smaller volume standard deviation and slower speech rate. Raters judged the left-frontal patients to be the most impaired group followed by the right-frontal and right-subcortical patients. Acoustical analysis revealed that left-frontal patients had significant difficulty with duration cues including slower speech rates and more silence in their samples than normal controls.

providing some support for the cue-lateralization hypothesis. Although not statistically significant, right-frontal and right-subcortical patients used less pitch variation in their speech, consistent with findings in the literature regarding the ability to modulate pitch range. Left subcortical patients did not demonstrate prosodic impairment. In addition, analyses suggested that men and women used pitch and duration cues differently to express prosodic emotion and that women had a happy tone whereas men an angry tone when speaking about a neutral topic. These findings support the notion that fronto-striatal circuitry, not just subcortical structures, is critical for prosodic emotion production.

Acknowledgements

Personally, I am grateful to my supervising committee who helped bring this work to fruition. My committee members offered their unique perspectives on this work which greatly added to the richness of the process. Howard Ehrlichman's contribution to the methods during the pilot phase was invaluable as was his support in improving the clarity and readability of the Results and Discussion sections. Nancy Foldi's experience with pragmatics and acoustical analysis helped me expand my perspective on how to interpret and discuss these data. I also very much appreciate the time and energy that Drs. Ronald Bloom and Candace Martin devoted as readers on this project. Their individual contributions have added greatly to the effort to publish aspects of this work. The support I have received from my committee chair, Joan Borod, goes well beyond the scope of this project. Joan has supported me since I arrived at graduate school with work in her lab, advice, and counsel about my future, and has provided emotional support to me personally when times were rough. I have greatly appreciated her unquestionable allegiance to me and I thank her for her presence.

I would also like to acknowledge three other people who significantly contributed to this project. Jennifer Speilman, M.A., graciously gave her time tutoring me regarding the nuisances of acoustical analysis. Not only did she direct me to the analysis software package used for this study, she spent time listening to samples and verifying analyses I conducted. Jennifer Sheth worked hand-in-hand with me through most every phase of this time-consuming project. Without the hundreds of hours that Jennifer spent working this project, it may never have been finished. Finally, I would like to thank Elizabeth Canino, Ph.D., whose contributions as both a friend and a colleague have been invaluable

directly to this project and to me personally. I appreciate her very much and the contributions she continues to make to my life.

The project was made possible by a grant from the National Institute of Mental Health (MH421712) to Dr. Joan C. Borod at Queens College.

Table of Contents

Abstract	iii
Acknowledgements	v
Table of Contents	vii
List of Tables	x
List of Figures	xi
Introduction	1
Background	1
Defining Emotion	3
Models of Emotion	5
Prosody	12
Definition	12
Methods to Study Affective Prosodic Expression	12
Anatomy of Prosody	15
Fronto-striatal-thalamic circuits	15
Amygdala PAG vocal pathway	16
Motor Cortex pathway	17
Literature review affective prosody production	17
Overview of prosodic processing literature	17
Perception	18
Cortical mediation of perceptual prosody	18
Subcortical mediation of perceptual prosody	19
Perceptual errors – analysis of specific acoustic cues based on subject errors	20
Production	22
Cortical mediation of posed expressive prosody	22
Subcortical mediation of posed expressive prosody	24
Acoustical cue production by brain-damaged subjects	25
Relationship of output acoustics and perceptual error	27
Spontaneous production studies	29
Current study	34
Methods	35
Participants	35
Overview	35
Subjects	35
Power Analysis	42
Raters	43
Collection of Expressive Data from Subjects	44
Experimental Tasks	44
Nonemotional Control Tasks	45
Experimental Procedures	46
Overview	46
Monologue Sampling	47
Monologue sampling procedure	48
Categories	48
Criteria for Categorization	49
Conferencing	50
Digital processing of speech samples	50

Rating Procedures	52
Measurement Introduction	52
Rating Scales	52
Rater Aggregation:	52
Rating Session Procedure	53
Acoustical Analysis	55
Acoustical parameters	55
Transformations	56
Data Preparation	56
Monologue Sampling Procedure	56
Ratings	57
Data Analyses	58
Introduction	58
Monologue Sampling Reliability	59
Self-training Trial Data	59
Rating Reliability	59
Demographic influence on intensity and confidence ratings	59
Subject Perspective - Rating Data	59
Subject Perspective - Acoustic Data	61
Rater Perspective - Acoustic to Category Relationships	61
Results	63
Monologue Sampling Reliability	63
Practice Trial Data	63
Rating Reliability	64
Demographic influence on intensity and confidence ratings	65
Subjects' Perspective - Rating Data	65
Correlations among Identification, Confidence, and Intensity Ratings	65
Identification Ratings	67
Confidence ratings	72
Intensity ratings	75
Subjects' Perspective - Acoustics	75
Introduction	75
Pitch mean	75
Pitch standard deviation	76
Pitch slope	77
Pitch regression 1 st half sample	78
Pitch regression 2 nd half sample	79
Pitch regression entire sample	79
dB standard deviation	80
Short Pause Rate	82
Long Pause Rate	82
Raters' Perspective - Acoustic to Category relationships	83
Introduction	83
Category	84
Introduction	84
Pitch mean	84
Pitch standard deviation	84
Pitch slope	85

Pitch regression 1 st half sample.....	85
Pitch regression 2 nd half sample.....	85
Pitch regression entire sample	85
dB standard deviation	86
Speech rate	87
Percent of sample with no voicing.....	87
Short Pauses	88
Long Pauses	89
Intensity.....	89
Introduction.....	89
Angry model	89
Happy model.....	89
Sad model.....	90
Discussion.....	91
Emotion Profiles	94
Gender specific emotion profiles.....	96
Women sound happy and men sound angry	98
Emotional communication hypotheses	101
Limitations	105
Future Directions	106
Conclusions.....	107
Appendix A.....	110
Appendix B.....	120
Appendix C	128
References.....	131

List of Tables

Table 1: Lesion location of subjects	37
Table 2: Demographics by Gender	39
Table 3: Demographics	39
Table 4: Ethnicity	39
Table 5: Subject Demographics	40
Table 6: Voice fundamental frequency/low-pass filter level conversion table.....	51
Table 7: Confidence and Intensity scales.....	53
Table 8: Examples of “Category” and “Strength” variable creation	57
Table 9: Emotion by Group monologue estimated mean agreement rates.	64
Table 10: Multiple Regression – Demographic Prediction of Intensity Ratings	66
Table 11: Multiple Regression – Demographic Prediction of Confidence Ratings.....	67
Table 12: Rating data descriptive statistics broken down by Gender and Group.....	68
Table 13: Group x Emotion effects - Best Sample Pitch Acoustics	86
Table 14: Group x Emotion effects - Best Sample Volume and Duration Acoustics.....	87
Table 15: Gender x Emotion effects - Best Sample Acoustics	88
Table 16: Mixed-effects model results - Acoustical parameter predictors of intensity....	91

List of Figures

Figure 1: Group x Emotion Identification Ratings	69
Figure 2: Group x Category Frequency Counts	70
Figure 3: Lesion site x Emotion Identification Ratings	71
Figure 4: Gender x Emotion Identification Ratings	72
Figure 5: Gender x Category Frequency Counts	73
Figure 6: Gender x Emotion Confidence Ratings	74
Figure 7: Pitch Standard Deviation by Lesion Site	77
Figure 8: Long Pause Rate by Group	83

Introduction

Background

The study of functional specificity of neuroanatomical structures originates from the largely discounted theory of phrenology proposed by Franz Gall, which suggests that different parts of the brain are responsible for specific functions depending on the bumps present on the skull. Systematic study of functional localization and hemispheric specialization was moved beneath the bones of our skull to the brain itself by Bouillaud (1825) who suggested that damage to the frontal lobes resulted in a loss of speech, and later, by Broca (1861) further specifying that damage to just the left frontal lobe resulted in impairment (Harrington, 1998). The study of localization of emotions began soon after with the case study of Phineas Gage and the seminal studies by Hughlings Jackson (1874; 1880) which were among the first to study the neurological underpinnings of emotion (Davidson, 1998).

The examination of emotional mechanisms included subcortical structures beginning in the 1900's. Bard (1929) proposed that the hypothalamus plays a critical role in emotions based on findings showing that electrical stimulation of the hypothalamus produces a sustained sympathetic arousal. Papez (1937) proposed a more complicated anatomical model of emotion suggesting that different parts of the brain may mediate different components of emotions. Papez proposed that the hypothalamus, thalamus, cingulate gyrus, hippocampus, and related interconnections form a loop mediating emotional expression, while the hypothalamic component of the circuit is involved in emotional expression. This model was further developed and named the "limbic system" by McLean (1949) (Gainotti, 2001). More recently, different models have been advanced based on data suggesting that both the cortical/subcortical dichotomy and also the

right/left hemisphere distinction may be relevant to understanding the cerebral organization of emotions (Borod, 1992). By investigating the relationship between brain and behavior, we can independently inform models of neural organization as well as models of cognitive processing (Baum & Pell, 1999).

Emotional prosody refers to suprasegmental features such pitch, loudness, timbre, tempo, stress, accent, pauses, intonation and melody that are used to convey emotion vocally (Ross, 1993). Studies have examined posed (Baum & Pell, 1997; Bloom, Borod, Obler, & Koff, 1990; Borod & Koff, 1990; Bradvik, Dravins, Holtas, Rosen, & al., 1991; Cancelliere & Kertesz, 1990; Gorelick & Ross, 1987; Martin, Borod, Alpert, Brozgold, & Welkowitz, 1990; Shapiro & Danly, 1985; Trauner, Ballantyne, Friedland, & Chase, 1996; Tucker, Watson, & Heilman, 1977) and spontaneous prosodic expression in brain-damaged subjects (Borod, Koff, Lorch, & Nicholas, 1985; Gorelick & Ross, 1987; Weintraub, Mesulam, & Kramer, 1981)[nonemotional prosody]). Studies of acoustics have examined the vocal properties of speech (Balan & Gandour, 1999; Dykstra, Gandour, & Stark, 1995; Hird & Kirsner, 1993; Ouellette & Baum, 1994; Ryalls, Joannette, & Feldman, 1987), and acoustical properties produced during emotional speech by brain-damaged speakers (Gandour, Larsen, Dechongkit, Ponglorpisit, & Khunadorn, 1995)[posed – happy, sad, neutral]; (Baum & Pell, 1997)[affective prosody – posed/spontaneous]; (Schirmer, Alter, Kotz, & Friederici, 2001)[posed]; (Pell, 1999c)[posed]; (Pell, 1999a). Yet, the lateralization of neuroanatomical mechanisms supporting emotional prosody remains unclear, and researchers have yet to fully characterize the ways in which speech acoustics provide cues to emotional states (Bachorowski, 1999). Furthering our understanding of the underlying neuroanatomical mechanisms of emotional prosodic expression and its relationship to different acoustical

parameters of basic emotions could have several implications for the rehabilitation of patients who have lost the ability to convey emotion prosodically. Specifically, this could improve health-care providers' predictions of at-risk patients, speed diagnosis of communication problems, facilitate retraining, and limit the negative impact of social and behavioral effects of misdiagnosis.

This study examined emotional spontaneous speech produced by patients with unilateral braindamage and neurologically intact matched controls. The subjects' spontaneous speech was sampled and judged by naïve raters for emotional intensity, emotional identification and confidence of their emotional identification rating. Research in vocal expression has demonstrated that listeners (decoders) reliably and accurately perceive emotions in the voices of human subjects (encoders) at rates that are six-fold better than chance (Sobin & Alpert, 1999) and are generally better than "objective" algorithms at detecting emotion (Rabinov, Kreiman, Gerratt, & Bielałowicz, 1995). Additionally, each speech sample underwent acoustical analysis to measure pitch, speech rate, and amplitude properties. Several issues were addressed in the experiment including the examination of the anatomical localization of emotional prosody production across different brain-damaged groups; the relationship between specific acoustical parameters and normal raters' judgments of the emotional production of brain-damaged patients; and methodological issues in the current literature applied to the study of brain-damaged populations.

Defining Emotion

The construct of emotion is difficult to operationalize, in part, due to the complexity of the phenomenon, and is further complicated by the numerous approaches available to emotion researchers. Broadly speaking, psychological, neurobiological, and

neuropsychological theories of emotion frequently address different aspects of emotion. Many controversies exist within emotion psychology because of the failure to distinguish among different classes of emotional phenomena (Scherer, 2000) which is exacerbated by the numerous acceptable ways to categorize emotion. The plethora of emotion research available, over 8500 published citations in the last five years alone, complicates these controversies (Cacioppo & Gardner, 1999). According to Heilman (1998), the domain of emotion can be divided into two major divisions: emotional experience and emotional behavior. Emotional experience can be momentary (e.g., happy, sad) or a mood (e.g., depression, anxiety) defined as a prevailing emotional experience over time. Emotional behaviors include autonomic-endocrine-visceral responses and emotional communication (verbal-semantic, facial, prosodic, and gestural-postural) with receptive and expressive components (Heilman & Gilmore, 1998).

Scherer (2001) differentiates five distinct affective states which vary on integral characteristics, including intensity, duration, synchronization, event focus, appraisal elicitation, rapidity of change, and behavioral impact. Scherer (2001) also distinguishes moods, interpersonal stances, attitudes, and personality traits from emotion, which he defines as episodes of coordinated changes in several components in response to external or internal events of major significance to the organism (e.g., anger, sadness, joy, fear). Moods, though closely related to emotions, differ in that moods are seen as diffuse states most pronounced as a change in subjective feeling, of low intensity and long duration, often without apparent cause (e.g., cheerful, gloomy, irritable, depressed). In contrast, interpersonal stances are affective positions taken toward another person in a specific interaction (e.g., distant, cold, supportive), attitudes are relatively enduring beliefs or preferences (e.g., hating, desiring), and personality traits are stable behavioral tendencies

typical for a person (e.g., nervous, reckless, jealous). This conceptualization differentiates the many different “feelings” that one might experience, but is incomplete in the sense that it does not operationalize how those “feelings” might be put into action.

As noted, Heilman (1998) conceives two major divisions within emotion: emotional experience and emotional behavior. Similar to Scherer (2001), emotional experience can be momentary emotions (e.g., happy, sad) or moods (depression, anxiety). Supporting this definition of emotion, many emotion psychologists stress the episodic nature of emotion and take the position that an emotional episode requires a noticeable change in functioning brought about by some external or internal triggering event lasting for a certain duration, decreasing in intensity until it fades away. Heilman’s (1998) emotional behaviors include autonomic-endocrine-visceral responses that represent the transformation of emotional experience into adaptive behaviors and emotional communication systems (verbal-semantic, facial, prosodic, and gestural-postural). These convey and receive messages of intent to further the likelihood that behavioral responses will be successful for the organism. Heilman’s categorization of emotion, therefore, not only allows us to clearly define what we are studying, but it also gives us a tool to categorize current models of emotion and a lens to examine the vast literature available.

Models of Emotion

Classes Models of emotion can be classified into at least three distinct classes, including dimensional models, discrete emotion models, and componential models (Scherer, 2001). Briefly, dimensional models can be either unidimensional or multidimensional in scope. Unidimensional models are based on the assumption that one dimension such as activation/arousal or emotional valence (e.g., positive/negative) is sufficient to make important analytic distinctions. Multidimensional models use several

dimensions to explain emotional experience and behavior and have been the basis of recent physiological and neuropsychological emotion research (e.g., Davidson, 1992; Davidson, 1993; Lang, Greenwald, Bradley, & Hamm, 1993). Alternately, discrete emotion models stress that during evolution a number of fundamental emotions have developed with distinct neural circuitry and clear behavioral outputs. In contrast, basic emotion models propose that fundamental emotions such as anger, fear, joy, sadness, or disgust have evolved into a number of adaptive emotional strategies, each having its own specific eliciting conditions, specific physiology, expressive, and behavioral reaction patterns.

Lastly, componential models propose that emotions are elicited by controlled and uncontrolled cognitive evaluation of antecedent situations and events and that the patterning of the reactions in the different response domains (physiology, expression, action tendencies, and feeling) is determined by the outcome of this evaluation process. One of the major features of componential theories is the effort to render the link between the elicitation of emotion and the response patterning more explicit, differentiating these models from others that generally ascribe to fixed behavioral patterns rather than dynamic patterns. Modern theorists who ascribe to componential processes of emotion acknowledge the existence of an array of components involved in emotional processing (e.g., processing modes) which can be studied in isolation or together to determine their relationship to each other (e.g., Borod, 1993b; Borod et al., 1985; Bowers, Bauer, & Heilman, 1993; Gainotti, Caltagirone, & Zoccolotti, 1993; Semenza, Pasini, Zettin, Tonin, & Portolan, 1986; Starkstein, Federoff, Price, Leiguarda, & Robinson, 1994).

Models Frequently, discrete and basic emotion models directly address momentary emotional experiences of the organism. A recent discrete emotion model has

been proposed by Panksepp (Panksepp, 1992, 1998) outlining seven emotional systems that are neuroanatomically, neurophysiologically, and neurochemically distinct including "seeking" (e.g., anticipation), "rage" (e.g., anger and frustration), "fear" (e.g., flight), "panic" (e.g., separation distress), "lust" (e.g., sexual desire), "care" (e.g., maternal nurturance and social attachment), and "play" (e.g., social engagement). Each of these neural circuits has very clear behavioral outputs, but it is expected that various interactions among these systems can lead to second order emotive states consisting of blended activities across the primary systems.

Alternately, Plutchik (1962; 1980) proposed a basic emotions model according to fundamental phylogenetically continuous classes of motivation identified by ethological research. Consistent with this, Ekman (1973; 1992; 1994; 1980; 1972) has shown strong evidence for the universality of discrete emotion patterns, particularly in emotional facial expressions which can be accurately identified across ethnic groups and cultures. Furthermore, basic emotion theorists postulate mechanisms of emotion mixing to explain the wide variety of emotional states generated from a small number of 7 to 14 basic emotions. Together, discrete and basic emotion theories propose the basic building blocks of momentary emotional experience.

Emotional behaviors include autonomic-endocrine-visceral responses and emotional communication systems with both expressive and perceptive aspects. Several levels of analysis are possible: analysis may involve specific emotional behaviors in isolation, relationships among emotional behaviors, or the relationships between emotional behaviors and emotional experience systems. Componential models of emotion are well equipped to conduct these analyses and are currently being utilized to assess whether there is a "central processor" of emotions or if emotion components are

subserved by different parts of the brain (Gainotti, 2001). For example, Borod (Borod, 1993b; Borod et al., 2000) describes a componential approach to emotion that examines whether processing mode, communication channel, emotional dimensions, and discrete emotion are mediated by a general processor or by separate processors.

Neuropsychological Theories Current neuropsychological theories regarding emotional processing are rooted in early clinical observations that distinguish emotional sequelae following right versus left hemisphere brain damage. In general, it was observed that subsequent to right hemisphere damage, many patients exhibit a generalized flattened affect with a concomitant monotonous speech production pattern, whereas subsequent to left hemisphere damage patients appear to produce relatively normal intonation and retain the ability to produce appropriate melody and rhythm for singing (Jackson, 1915). These clinical observations were dissociable from the observations of Broca (1861) regarding left hemisphere lateralization for language functioning and led Hughlings Jackson (1915) to theorize that language processing centers were contained within the left hemisphere while emotional centers were within the right hemisphere (Baum & Pell, 1999).

Current proposals include the Right-hemisphere hypothesis, which maintains that the right hemisphere is dominant for emotional processing (Borod, Zgaljardic, Tabert, & Koff, 2001; Borod, 1992; Borod & Madigan, 2000; Bryden & Ley, 1983; Buck, 1984; Heilman, Blonder, Bowers, & Crucian, 2000; Ross, Thompson, & Yenkosky, 1997). This is supported at the behavioral level according to Borod (Borod, 1992), as emotional processing involves strategies (e.g., nonverbal, synthetic, integrative, holistic, and Gestalt) and functions (e.g., pattern perception, visuospatial organization, and visual imaging) for which the RH is superior. Tucker (1981; 1989) also suggested that two

cognitive features favor the right, rather than the left, hemisphere in the processing of emotional information including the tendency of the right hemisphere to represent experience in an analogical (not verbal) code and its ability to integrate information in a synthetic and holistic manner rather than treating it in a sequential, analytical fashion. According to Borod et al. (1998), at the neuroanatomical level, the right hemisphere has been described as having greater multimodal integration (Goldberg & Costa, 1981; Semmes, 1968), interlobular organization (Egelko et al., 1988), interregional connectivity (Gur et al., 1980; Thatcher, Krause, & Hrybyk, 1986; Tucker, Roth, & Bair, 1986), stimulus-evoked physiological activity (Trotman & Hammond, 1989), and horizontal axonal connectivity (Springer & Deutsch, 1998; Woodward, 1988) than the LH. Thus, the RH may be particularly suited to the complex, multimodal, and integrated nature of emotional processing.

The valence hypothesis is an alternate proposal that has two variations (Borod, 1992). One postulates right hemisphere dominance for negative emotions and left hemisphere dominance for positive emotions, across processing modes (Heller, 1990; Sackeim et al., 1982; Silberman & Weingartner, 1986). The other version postulates that there is differential hemispheric specialization as a function of valence for emotional expression and experience, whereas the right hemisphere is dominant for emotional perception across valence (Bryden, 1982; Davidson, 1984). Differential specialization occurs for frontal regions, and right hemisphere dominance pertains to posterior regions (Davidson, 1984). It is believed that because negative emotions are linked with survival (e.g., removing oneself from danger), a system is required that is sensitive to multimodal inputs that can quickly scan the environment and evaluate the situation (Borod, Koff, & Buck, 1986). Such behaviors involve more Gestalt, synthetic processing (RH functions)

than discrete, focused analysis (LH functions). Positive emotions, by contrast, may be more dominated by LH functions, being more linguistic and communicative than emotional and reactive (Borod et al., 1986; Borod, Caron, & Koff, 1981). Interestingly, it has also been suggested that social emotions (e.g., affection) are mediated by the LH and primary emotions (e.g., sadness) by the RH; see studies by (Langer, Pettigrew, Wilson, & Blonder, 1998; Ostrove, Simpson, & Gardner, 1990; Ross, Homan, & Buck, 1994). This notion is supported by the observation that the operation of social display rules (presumably left hemisphere-controlled) that normally inhibit the expression of primary emotions (believed to be right hemisphere-mediated) (e.g., (Buck & Duffy, 1980; Ross et al., 1994)).

The Motoric Direction Hypothesis (Borod et al., 1998) attempts to incorporate motoric action with other components of emotion, hypothesizing an "approach system" and "withdrawal system" (Davidson, Abercrombie, Nitschke, & Putnam, 1999). The "approach system" facilitates appetitive behavior and generates certain types of positive affect that are approach-related (e.g., enthusiasm). The "withdrawal system" facilitates removal of an individual from sources of aversive stimulation and generates certain forms of negative affect that are withdrawal-related. Davidson (1999) proposes that the "approach system" and positive affect are mediated by left prefrontal cortex while the "withdrawal systems" and negative affect are supported by right prefrontal cortex, which finds some support in lesion and neuroimaging data. Further, the dorsolateral, ventromedial, and orbitofrontal sectors of the prefrontal cortex (PFC) mediate different aspects of emotion systems. Specifically, the ventromedial PFC is critical to the anticipation of future positive and negative consequences of actions and is involved in the representation of elementary positive and negative states in the absence of immediate

elicitors. The dorsolateral PFC mediates the representation of the goal states to which these affective states are directed. This model also discusses the relevance of the amygdala, ventral striatum, anterior cingulate and insular cortex in emotion, the specifics of which are beyond the scope of this review.

Also recently, Rolls (2000) proposed the Reinforcement Theory of Emotion. Rolls' theory does not address potential cortical asymmetries in emotional processing, but attempts to explain the interaction of emotional processing and emotive action systems with basic motivational and reinforcement systems, naming specific cortical and subcortical structures mediating these interactions. This theory defines emotions in terms of behavioral states experienced by organisms in response to changing reinforcement contingencies and proposes neural mechanisms for the relationship between emotional encoding and behavior responses (Rolls, 2000). According to this theory, emotional encoding requires the amygdala and orbitofrontal cortices to assess primary reinforcers received from each sensory system (e.g., pleasant or painful touch, or visually perceived emotional facial expressions). Extensive connections of these structures to the periaqueductal grey (PAG) trigger instinctive, hardwired behavioral reactions that support survival. Direct stimulation of topographically organized PAG locations have been shown to produce fight / flight behaviors which puts it in an ideal position to mediate responses to the emotional experiences of the organism (Bandler & Shipley, 1994). Conscious appraisal of emotional events is facilitated further by the retrieval and creation of associations between primary reinforcers and neutral stimuli by the amygdala and hippocampus which are routed to the orbitofrontal cortex, other prefrontal areas, and language systems.

As stated earlier, there is value in analyzing specific emotional components in isolation, e.g., to build toward or find support for specific theories of emotion. The success of emotional behaviors is facilitated by the expression and perception of emotion through emotional communication systems including facial, verbal, prosodic, and gestural. Expressive prosody, one component of these systems, will now be discussed in detail.

Prosody

Definition Affective prosody can be defined as suprasegmental features such as pitch, loudness, timber, tempo, stress, accent, pauses, intonation and melody that are used to convey emotion vocally (Ross, 1993). According to Baum and Pell (1999), there are currently 3 major hypotheses regarding the neuroanatomical regions active in affective prosodic processing. The first hypothesis contends that all aspects of affective prosody are processed in the right hemisphere, whereas linguistic prosody is localized in left hemisphere. Another view posits that both the comprehension and production of prosody are subserved largely by subcortical regions and are not lateralized to one or another hemisphere. Lastly, several investigations support the theory that individual prosodic features (e.g., fundamental frequency, syllable duration) may be independently lateralized (Van Lancker & Sidtis, 1992).

Methods to Study Affective Prosodic Expression Evaluating the prosodic abilities of brain-damaged patients is the primary tool available to researchers studying the neural mechanisms of affective prosody (Baum & Pell, 1999). This has been more true for prosodic production than perception. Perception studies have used other techniques instead of or in conjunction with using brain-damaged patients such as lateralized stimulus presentation in normal subjects (e.g., Blumstein & Cooper, 1974; Zurif &

Mendelsohn, 1972), *and* PET (e.g., George et al., 1996; Zatorre, Evans, Meyer, & Gjedde, 1992), fMRI (e.g., Buchanan et al., 2000), and ERP (e.g., Twist, Squires, Spielholz, & Silverglide, 1991) and functional imaging techniques (Twist et al., 1991)[ERP with BDs]. Perceptual studies of emotional attributes of speech have contributed greatly to the hypothesis that right hemisphere mechanisms are selectively engaged in the processing of affective prosody (Baum & Pell, 1999).

Affective prosody production can be elicited with more naturalistic, “spontaneous” conditions, as well as with more formal, “posed” elicitation procedures. Posed procedures are designed to elicit planned, voluntary behavior, whereas spontaneous conditions presumably tap involuntary emotional expression (Borod, 1993a). In posed conditions, subjects are requested to say a neutral-content sentence in an emotional tone of voice (oral command) or to listen to a model intoning a sentence and then to repeat the sentence with the same emotional tone (imitation). Spontaneous speech can be elicited under a variety of conditions. During an unstructured narrative, the subject is prompted with open-ended questions by a skilled interviewer which can produce long fluent stretches of speech for analysis. Another option is instruction monologues, in which the subject is asked to instruct an imagined listener to perform a task, such as constructing a model of a house. Finally, researchers can use an extended descriptive narrative where the subject retells a story of a past experience: this is more controlled than an unstructured narrative that includes the interaction with the interviewer (Beckman, 1997).

When these techniques are used to evaluate affective prosody, neutral nonemotional materials are often used for control purposes (e.g., with aphasics) (Borod, 1993a). Subjects’ speech samples are audio recorded and later evaluated by judges for

accuracy, expressivity, or intensity, allowing differences among the patient groups to be compared. The speech can also be subjected to computerized procedures involving speech filtering (e.g., (Blumstein & Cooper, 1974)) or acoustical analysis (e.g., (Alpert, Rosen, Welkowitz, Sobin, & Borod, 1989; Welkowitz, Bond, & Zelano, 1990)) (Borod, 1993a). The three primary acoustical parameters used to assess prosodic functioning are fundamental frequency (pitch), duration (speech rate), and amplitude (loudness) (Baum & Pell, 1999).

Fundamental frequency (F_0) is the rate at which the vocal folds vibrate and is typically measured in hertz (Pittam & Scherer, 1993). Normal speech usually contains a range of fundamental frequencies as a restricted F_0 range is perceived as flat or monotonous (Pell, 1999b; Pittam & Scherer, 1993). In general, F_0 varies continuously and often rapidly, but does so around a base level for each individual (Pittam & Scherer, 1993). The F_0 mean and standard deviation are common measures used to quantify the emotionality of speech (Banse & Scherer, 1996). Perturbations in F_0 manifested as fast random fluctuations in period duration are known as jitter (Pittam & Scherer, 1993), which is also used to quantify emotional changes in speech (Bachorowski & Owren, 1995). Duration measures can evaluate periods of articulation and silence, and quantify the different kinds of silence, e.g., pauses between words versus hesitation (Pittam & Scherer, 1993). The mean and standard deviation of amplitude measures reflect volume level and changes in speech. Like F_0 , the short-term perturbations of amplitude can be measured to reflect emotional changes in speech (Bachorowski & Owren, 1995). Studies using acoustical analyses, together with studies using independent judgments of emotional intensity and accuracy of brain-damaged speech production, have provided

insight into the underlying neural mechanisms and hemispheric specialization of affective prosodic expression.

Anatomy of Prosody

Fronto-striatal-thalamic circuits Prefrontal cortices give rise to two separate pathways that together produce speech containing both linguistic and affective information (Holstege & Ehling, 1996; Jurgens, 1994; Jurgens & Zwirner, 1996). Five parallel anatomic circuits originating in the motor (Broca's areas 6, 4, and 1), oculomotor (Broca's area 8), dorsolateral (Broca's areas 9, 10), cingulate (Broca's area 24), and orbitofrontal (Broca's 10, 11) cortices link regions of the frontal cortex to the striatum, globus pallidus, substantia nigra, and thalamus, creating distinct functional units (Cummings, 1995). All but the oculomotor circuit influence vocal output through a series of subcortical connections.

The orbitofrontal circuit is particularly important to emotion and can be divided into a medial and lateral component. Damage to the medial orbitofrontal cortex is associated with antisocial, impulsive behaviors and abnormal responses to social stimuli while damage to the lateral circuit is associated with irritability, emotional incontinence, and tactlessness (Mega & Cummings, 1994). The medial circuit projects to the ventral striatum, including the nucleus accumbens giving the medial circuit direct access to reinforcement circuits while the lateral circuit sends its projections to the ventral medial caudate. The cingulate circuit also projects to the ventral striatum and damage to its circuitry is associated with apathy, akinetic mutism, indifference to pain, and poverty of speech (Mega & Cummings, 1994). The dorsolateral circuit is important for cognitive functions and sends projections to the dorsolateral caudate. Damage to this circuit is associated with organizational difficulties, problems implementing memory search

strategies, and impaired set shifting. Each of these circuits continue to send topographically separated afferents to the globus pallidus and substantia nigra followed by the thalamus and back to its origin, completing the loop. The medial OFC, lateral OFC and cingulate circuits are not interconnected, but each has reciprocal connections with dorsolateral circuit and working memory areas (Kaufer & Lewis, 1996). Several studies have found evidence for the importance of prefrontal cortex in prosody or emotion including lesion studies (Breitenstein, Daum, & Ackermann, 1998; Shapiro & Danly, 1985; Starkstein et al., 1994), case-studies (Dykstra et al., 1995; George et al., 1996), and imaging studies (Buchanan et al., 2000; Lane, Reiman, Ahern, Schwartz, & Davidson, 1997; Lane et al., 1998; R. D. Lane et al., 1997; Morris, Scott, & Dojan, 1999; Pihan, Altenmuller, Hertrich, & Ackermann, 2000).

Amygdala PAG vocal pathway The cingulate and both orbitofrontal circuits project to the PAG directly and indirectly via the amygdala complex (Jurgens, 1994; Van Bockstaele, Bajic, Proudfit, & Valentino, 2001). The PAG also receives important inputs from the medial and lateral hypothalamus, and all inputs are topographically organized. When stimulated, the PAG produces important emotional behavior responses, including affective coordinated defense, flight responses, and quiescence coordinating blood flow, tachycardia, and movements (Bandler & Shipley, 1994; Fisk & Wyss, 2000). The PAG sends projections to mouth opening muscles, lower mouth muscles and tongue muscles via the caudal medullary lateral tegmental field and soft palate muscles, pharynx muscles, larynx muscles, intercostals muscles, and abdominal muscles via the nucleus retroambigous to produce vocalization (Farkas, Jansen, & Loewy, 1997; Jurgens, 1994). Complete mutism can be produced solely after PAG lesions (Esposito, Demeurisse, Alberti, & Fabbro, 1999).

Motor Cortex pathway The primary motor cortex circuit receives input from the cingulate and orbitofrontal circuits via connections with the dorsolateral circuit and also receives input to corticobulbar and corticospinal neurons from Broca's area, which is important for language production (Holstege & Ehling, 1996; Kaufer & Lewis, 1996). The motor cortex sends fibers via corticobulbar and corticospinal fibers to mouth opening muscles, peri-oral muscles, pharynx muscles, and tongue muscles via direct connections and indirect connections by way of premotor interneurons with motoneurons producing vowels and syllables necessary for language. In sum, these two pathways allow emotional, cognitive, memory, and motivational systems to influence and direct vocalization output and can be coordinated with emotional behaviors mediated via the PAG.

Literature review affective prosody production

Overview of prosodic processing literature The existing prosody literature has yet to provide a clear picture of the neural mechanisms controlling prosody (Baum & Pell, 1999) as there is some support for right-hemisphere dominance of both linguistic and affective prosody (e.g., (Starkstein et al., 1994)), linguistic prosody lateralized exclusively to the left hemisphere with affective prosody to the right (e.g., (Van Lancker, 1980)), and evidence that suggests bilateral control of prosody (e.g., (Dykstra et al., 1995)). It has also been proposed that subcortical structures (e.g., (Cancelliere & Kertesz, 1990)), rather than cortical structures, are essential for prosodic processing and, finally, that specific acoustic parameters may be selectively processed by either the left (e.g., duration) or right (e.g., fundamental frequency) hemisphere (e.g., (Van Lancker & Sidtis, 1992)). An extensive review of the prosodic expression literature, which is the primary focus of this study, is preceded by a brief review of the perception literature

which highlights the evidence for these opposing hypotheses and informs the broader goals of this proposal.

Perception

Cortical mediation of perceptual prosody Several prosodic perception studies support the hypothesis that prosody is primarily mediated cortically. Heilman (1984) studied patients with right- or left-hemisphere damage and control subjects to determine whether the RBD patients had global or limited prosodic defects. Compared to LBD patients and controls, RBD subjects had decreased comprehension of emotional prosody. Both LBD and RBD groups, which did not differ, had worse comprehension of propositional prosody than controls. The author suggested that the right hemisphere seems to be dominant for comprehending emotional prosody but not propositional prosody. Bryan (1989) also found that although RBDs were worse than controls on 13 linguistic prosodic perception tasks than control subjects, LBDs also performed worse than controls on 10 measures and RBDs were only worse than LBDs on 8 measures.

Similar findings were found more recently by Ross et al. (1997) who administered an aprosodia test battery to 10 LBDs and 12 RBDs over a 3-year period. Affective prosodic comprehension and discrimination were tested by presenting 3 conditions to subjects including presenting a sentence (e.g., "I am going to the other movies"), a monosyllabic "sentence" (e.g., "ba ba ba ba ba ba"), or asyllabic "sentence" (e.g., "aaaaaaahhhhhh") in neutral, happy, sad, surprised, angry, or disinterested intonation. In total, RBDs were worse than LBDs and controls on most measures, though LBDs were often more impaired than controls as well, and occasionally as impaired as RBDs. However, the authors make the point that reducing verbal-articulatory conditions robustly

improves the performance of left but not right brain-damaged patients, supporting the supposition that affective prosody is strongly lateralized to the right hemisphere.

Pell and Baum (1997) administered receptive linguistic and affective prosody tasks to 9 RBD, 10 LBD, and 10 age-matched control (NC) subjects. Neither patient group was impaired relative to normals in discriminating prosodic patterns or recognizing affective tone conveyed suprasegmentally. Although neither group appeared to display a receptive disturbance for emotional prosody, LBDs were less able to utilize the terminal rise and fall of the F_0 contour present in the linguistically defined categorical properties of prosodic stimuli than either NC or RBD subjects. LBDs had more difficulty on linguistic prosody tasks than RBDs and NCs even when semantic information biased the target response. Thus, the authors were able to show support for the theory that linguistic prosody is left lateralized, but were unable to show support for the right lateralization of affective prosody.

Pell (1998) administered 6 perception tasks that varied both functional (i.e., linguistic/emotional) and structural (i.e., acoustic) attributes of a common set of base stimuli to RBD, LBD, and NC subjects. Three tasks explored linguistic prosody elements, and 3 tasks examined the comprehension of emotional-prosodic meanings by the same listeners. Results revealed that both RBD and LBD subjects were impaired in interpreting the emotional meaning of prosodic contours, but that only LBD subjects displayed subnormal capacity to perceive linguistic specifications via prosodic cues. The performance of the RBD and LBD subjects was also selectively disturbed when certain acoustic properties of the stimuli were manipulated.

Subcortical mediation of perceptual prosody Several perception studies point to the possibility that subcortical structures are critical to prosodic processing. Starkstein et

al. (1994) tested 59 patients with acute stroke for the presence of comprehension aprosody. Patients with aprosodia for emotional comprehension also showed a higher frequency of right-hemisphere lesions involving the basal ganglia and the temporoparietal cortex and more severe frontal and diencephalic atrophy. Pell (1996) administered a linguistic and affective prosody perception battery to 11 patients affected by idiopathic Parkinson's disease (PD), and 11 NCs. PD patients showed a reduction in the ability to identify the linguistic- and affective-prosodic meaning of utterances relative to control subjects, without a concurrent loss in the ability to perceive phonemic stress contrasts or discriminate prosodic patterns. The authors hypothesized that the basal ganglia form part of a functional network dedicated to prosodic processing.

Breitenstein (1998) examined the differential contribution of cortical and subcortical brain structures in emotional processing by comparing 16 cortical RBDs (half anterior/half posterior), 16 cortical LBDs (half anterior/half posterior) and 14 Parkinson's disease patients with primarily subcortical dysregulation of the basal ganglia. Only patients in the more advanced stages of Parkinson's disease and patients with focal damage to the (right) frontal lobe differed significantly from controls in affective prosody recognition. In light of the extensive connections of the basal ganglia with frontal structures, these findings imply involvement of the fronto-striatal circuitry in emotional processing (Breitenstein et al., 1998).

Perceptual errors – analysis of specific acoustic cues based on subject errors

Finally, there are mixed data that suggest perception of specific acoustic parameters is lateralized. Van Lancker and Sidtis (1992) and Pell and Baum (1997) examined the relationship between the affective perceptual errors of brain damaged patients and specific acoustical parameters. Van Lancker and Sidtis (1992) had 24 LBD and 13 RBD

patients and 37 NC subjects perform a typical affective-prosodic listening task using 4 emotional types (happy, sad, angry, and surprised). To determine if the 2 brain-damaged groups were exhibiting a similar pattern of performance with respect to their use of acoustic cues, the 16 stimulus utterances were analyzed acoustically, and the results were incorporated into an analysis of subjects' errors. Both patient groups were less accurate than control subjects on both linguistic and affective perception tasks and did not significantly differ from each other. A discriminant function analysis using acoustic cues that incorporated subjects' misclassifications revealed that the F_0 variability was the best predictor of RBDs' perceptual errors while syllable duration variability was most predictive for LBDs' errors, suggesting that different acoustical cues are utilized by each hemisphere to perceive meaning. Specifically, these results suggest that right-hemisphere damaged patients did not appear to make use of F_0 variability, but relied on duration cues to make their affective classifications, while left-hemisphere damaged patients were able to use F_0 information but failed to use duration cues adequately.

Conversely, Pell and Baum (1997) acoustically analyzed errors from comprehension tasks of affective and linguistic prosody made by 10 LBDs, 9 RBDs, and 10 control subjects. An analysis of the errors made on these tasks by LBD and RBD subjects examined whether each clinical group relied on specific and potentially different acoustic features in comprehending prosodic stimuli. Although acoustic analyses of the utterances reaffirmed the importance of F_0 cues in signaling affective and linguistic prosody, analyses of subjects' affective misclassifications did not suggest that LBD and RBD patients were biased by different sets of acoustic features of prosody in judging their meaning, in contrast to Van Lancker and Sidtis (1992). In total, the perception literature simply does not elucidate a systematic neural organization of prosody, and

further study will be needed to determine the contributions of cortical and subcortical structures and the functional organization of prosodic acoustical cues (Baum & Pell, 1999)

Production

Cortical mediation of posed expressive prosody Fewer studies have been conducted on prosodic production relative to perception, but the same basic issues are paralleled in the production literature including evidence of the importance of cortical structures for prosody production. Tucker et al. (1977) conducted two experiments, the first examining the ability of 8 right brain damaged and 8-normal controls to repeat neutral sentences with specific emotional tones, which were then rated for emotional accuracy by three listeners. The second experiment compared 8 RBDs to 7 conduction aphasics with left hemisphere damage on their ability to discriminate differences in prosodic patterns. Results from both experiments show significant impairment in RBDs relative to the comparison group, lending support to the right hemisphere lateralization of prosody. However, neither experiment compared performance of LBDs, RBDs, and normals together -- the first omitting LBDs and the second omitting normals, making the interpretation of the results more complicated. LBDs were not error free in the perception condition and the absence of the LBDs in the expression experiment fails to rule out equal deficits of LBDs and RBDs in prosodic expression. Other linguistic prosody (e.g., (Bradvik et al., 1991; Weintraub et al., 1981)) and affective prosody studies (e.g., (Heilman, Scholes, & Watson, 1975)) also demonstrate prosodic perceptual deficits associated with right hemisphere damage but lack appropriate comparison groups.

Shapiro and Danly (1985) provide scientifically rigorous support for right hemisphere primacy of affective prosody. They studied sixteen right-handed adult males with localized insult to either the right or left hemisphere and five control subjects without brain damage: subjects read aloud target sentences embedded in paragraphs while intoning their voices in a declarative, interrogative, happy, or sad mode. An acoustical analysis of the speech wave was performed, and results revealed right-anterior (pre-Rolandic) and right-central (pre- and post-Rolandic) brain-damaged patients spoke with less pitch variation and restricted intonational range across emotional and nonemotional domains, whereas patients with right posterior (post-Rolandic) damage had exaggerated pitch variation and intonational range across both domains. No such deficits were found in patients with left posterior damage, whose prosody was similar to that of normal control subjects. It was suggested that damage to the right hemisphere alone may result in a primary disturbance of speech prosody that may be independent of the disturbances in affect often noted in right-brain-damaged populations.

Ross et al. (1988) examined the ability of five subjects with epilepsy undergoing a WADA procedure to repeat a sentence intoned in six different emotions following a model. In this test, a sodium amytal solution is injected into the internal carotid artery to temporarily 'deactivate' one hemisphere of the brain, allowing researchers to assess the functioning of one hemisphere in the absence of the other hemisphere. Acoustic parameters were measured from productions recorded prior to and during the Wada tests. Results revealed that mean F_0 and F_0 standard deviation were reduced during the test as compared to before the injection, suggesting a flattened affect as a result of deactivation of the right hemisphere (Ross et al., 1988). Ley and Bryden (1982) and Erhan et al. (1998) both showed support for right hemisphere dominance for prosodic processing by

using a dichotic listening paradigm to demonstrate left ear (right hemisphere) superiority for judging emotions and a right ear (left hemisphere) advantage for judging sentence content. Further support comes from Tompkins and Flowers (1985), Bowers et al. (1987), and Blonder et al. (1991), all of whom found significant prosodic perception deficits associated with RBDs relative to LBDs and NCs.

Subcortical mediation of posed expressive prosody However, some studies have failed to find prosodic deficits related to cortical structures, but instead find support for the importance of subcortical structures for prosody. In a recent study, Baum and Pell (1997) studied via acoustic analysis the ability of 7 RBD aphasic, 4 LBD aphasic, and age-matched control subjects (NC) to produce linguistic and affective prosodic contrasts at the sentence level via acoustic analysis. Their results demonstrated that both RBD and LBD subjects were able to appropriately utilize the acoustic parameters examined to differentiate both linguistic and affective sentence types in a manner comparable to NC speakers. Although not statistically significant, some irregularities in the global modulation of fundamental frequency and amplitude by RBD speakers were noted. Baum and Pell did not find support for specific engagement of the right hemisphere in the production of affective prosody. They suggested that prosody may not be functionally lateralized, but under some subcortical control and that particular acoustic parameters may be differentially lateralized.

Several studies do, in fact, support subcortical mediation of prosody production and the possibility that acoustic parameters are themselves lateralized. Cancelliere and Kertesz (1990) investigated the relationship between intrahemispheric location of lesion and disturbances of emotional expression and comprehension in 28 right-hemisphere and 18 left-hemisphere stroke patients. Patients were classified into aprosodic syndromes

according to test scores on measures including a standardized test of the expression, repetition, and communication of emotional prosody. The basal ganglia were most frequently involved in aprosodic syndromes followed by anterior temporal lobe and insula structures. The authors concluded that the basal ganglia are an important structure in the mediation of emotional expression and comprehension.

Bradvik (1991) tested the ability to perceive and express emotional and linguistic prosodic qualities of speech in 20 Swedish-speaking patients with right-sided brain infarcts and 18 normal controls. Patients showed disturbed abilities in the identification of emotional messages as well as in the identification and production of several linguistic prosodic qualities. Based on lesion analysis, their results supported not only the importance of the right hemisphere for identification and production of several linguistic and emotional prosodic qualities of speech, but also that speech prosody can be disrupted by purely subcortical lesions. The authors did not discuss whether subcortical lesions, focused only in the right hemisphere, were sufficient to produce prosodic deficits.

Acoustical cue production by brain-damaged subjects The theory that each hemisphere mediates different acoustic cues has been examined in the posed production literature. Ouellette et al. (1994) examined the acoustic correlates of spontaneous linguistic prosodic production of 9 LBD nonfluent aphasics, 8 RBD patients, and 9 normal speakers. Results indicated that RBD and LBD groups resembled normal speakers in the use of fundamental frequency and amplitude to encode stress, suggesting preserved abilities in both neurological populations. Pell (1999b) examined prosody production in 10 RBD individuals to explore potential acoustic differences between 10 normal speakers and those perceived to be emotionally "flat". Results indicate that RBD patients were significantly more likely to be perceived as lacking emotional inflection by

a group of normal listeners than age-matched normal speakers. Also, discriminant analysis showed that mean F_0 served as the principle cue to emotional contrasts of sad, angry, and happy. Furthermore, the subjects (e.g., five RBDs, one NC) obtaining low emotional ratings were considered separately and were shown to employ substantially significantly lower mean F_0 levels than their counterparts, explaining raters perception that these speakers were emotionally "flat". In sum, these results were seen to support a quantitative impairment for prosody in many RBD individuals.

In a separate study, Pell (1999a) administered a story completion task to matched groups of 10 RBD and 10 NC subjects. Subjects produced utterances that simultaneously specified 3 prosodic distinctions (emphatic stress, sentence modality, emotional tone); these were subjected to acoustic analysis to examine various F_0 attributes of the stimuli. Their results indicated that RBD speakers tended to produce fundamental frequency patterns that resembled normal productions in overall shape, but with significantly less F_0 variation. The RBD patients were also less reliable than normal speakers at transmitting emphasis or emotional contrasts when judged by the listener's.

Finally, in a third study, Pell (1999c) evaluated the acoustical timing characteristics produced by speakers with and without focal right-hemisphere damage. Utterances varying along different prosodic dimensions (emphasis, emotion) were elicited from each speaker using a story completion paradigm, and measures of utterance rate and vowel duration were computed. Results demonstrated many similarities in how RBD and healthy individuals encoded the temporal correlates of emphasis. However, the authors found significant differences in how RBD speakers employed temporal cues to specify some aspects of prosodic meaning, particularly for emotional content. Specifically, results revealed that RBDs produced sad utterances significantly faster than

normal subjects, which they believed led to poor discrimination between neutral, happy, sad, and angry samples. In conjunction with findings from previous studies, the authors suggest that RBD individuals have difficulty modulating vocal parameters (e.g., F_0 and duration cues) in the continuous manner required to allow differentiation of emotional inflection from neutral speaking. As such, the authors suggest RBDs are most disturbed when expressing prosodic representations that vary in a graded, rather than categorical, manner.

Relationship of output acoustics and perceptual error Finally, another means to evaluate the prosodic properties of affective speech output is to present prosodic production from a variety of emotional categories (e.g., happy, sad, angry) to naïve judges and request they identify the emotion being conveyed. Production samples can then be regrouped based on the judges' ratings of emotionality regardless of the target emotional category. An acoustical analysis of the regrouped samples allow researchers to ascertain an acoustical profile for specific emotions, identify relationships between accurately identified and misjudged prosodic output, and relate these findings back to the populations that produced the output.

For instance, in a study by Banse and Scherer (1996) 14 emotions varying in intensity and valence portrayed by professional actors were presented to 3 male and 9 female undergraduate decoders. Decoders were able to infer the vocally expressed emotions with much-better-than-chance accuracy and consistently found differences in emotions being recognized. Further, 224 portrayals were subjected to acoustic analysis to obtain profiles of vocal parameters for the different emotions. These suggested that the degree of intensity was typical for different emotions but also differentiate valence or quality aspects. Multiple regression revealed that "Happiness" was characterized by

average levels of F_0 mean energy and increases in speech rate relative to other emotions. In contrast, sadness was characterized by average levels of F_0 but was among the lowest in mean energy level and speech rate. Anger was characterized by high F_0 mean levels, high mean energy levels, and fast speech rate relative to other emotions.

Gandour, Larsen et al. (1995) examined the production of speech prosody in emotional contexts in Thai speaking patients with unilateral right hemisphere damage. Twelve RBDs and 9 normals read target sentences embedded in paragraphs that cued either a happy, sad, or neutral affect. Seven raters listened to the posed production and tried to identify the emotion being conveyed. Perceptual evaluations of their productions revealed a severe deficit in right hemisphere patients for both happy and sad utterances. Specifically, RBDs produced significantly fewer statements that could be accurately identified by raters; of those statements accurately identified, RBD statements were significantly longer in duration and narrower in energy variability than those produced by normals. Analysis of statements produced by RBDs and inaccurately identified by raters revealed that utterances were significantly longer in duration and higher in mean F_0 and energy than for on-target happy contexts. These results confirm that rater identification errors can be attributed to differences in acoustical cues and that RBDs produce significantly different acoustic patterns of affective prosody than normals.

Sobin and Alpert (1999) attempted to evaluate the acoustic cues that raters use to accurately identify emotional prosody. Thirty-five normal encoders practiced saying and then recorded five neutral versions of a standard sentence. Twenty emotion-inducing stories were then presented in a modified random block design. Following the reading of each story, encoders were asked whether an emotion had been experienced, and if so, which emotion (choices: fear, hate, anger, jealousy, sadness, guilt, joy) was experienced

as the primary and secondary emotional experiences, and to rate on a 10-pt scale the intensity of the primary emotion experienced, and the intensity with which the felt emotion was expressed. Decoders, specialized for a specific emotion, then rated target sentences for the presence of the emotion from “not present” to “extremely intense” on a 7-point scale. Acoustical analysis of the emotional sentences revealed that 8 of 12 acoustic variables were significantly different across emotions. Multiple regression analysis of decoder ratings and acoustic variables suggested that “Joy” was negatively associated with volume variance and positively associated with pitch variation, the number of emphasized syllables, and the duration of speech. On the other hand, “Sadness” was negatively associated with volume and pitch variance and also positively associated with the duration of speech. Finally, anger was positively associated with volume and pitch variance, while negatively associated with mean pitch and the number of emphasized syllables.

Spontaneous production studies Substantially less work has been done using spontaneous prosodic production compared to posed production studies. Based on this review, only 4 large-N studies have examined spontaneous affective prosody in addition to 5 case-studies. Of these case studies, 2 were not contributory, one did not evaluate any neuropsychological theories, and the other noted acoustic deficits but not anatomical relationships. Of the 3 that considered neuropsychological issues, different results emerged.

Ross and Mesulam (1979) describe a 39-year-old right-handed female and a 62-year-old left-handed male that both suffered prosodic deficits after a right-hemisphere stroke. Their observations led them to suggest that there are similar cortical areas subserving affective speech in the right hemisphere to the cortical areas in the left

hemisphere subserving prepositional speech. However, Lebrun et al. (1985) failed to find a prosodic deficit in a bilingual right handed male (age not reported) despite showing clear right hemisphere deficits following right temporal lobe and supramarginal gyrus resection for the extraction of a neoplasm.

Dykstra et al. (1995) clinically observed a spontaneous prosody deficit in a 20 year-old female with a history of seizures, but did not include spontaneous prosody as a focus of study. The authors did however quantify a posed affective prosody deficit showing that timing and intensity were impaired, while F_0 was relatively spared. Blonder (1995) evaluated spontaneous speech prosody in a 77 year-old woman, finding reduced and restricted F_0 post stroke affecting the right frontotemporo-parietal regions and the right basal ganglia; and Leentjens et al. (1998) assessed affective prosody in schizophrenics, but no neuropsychological theories of prosodic processing were evaluated.

The earliest of the four large-N studies (Borod et al. 1985) evaluated spontaneous emotional expression in 12 right-brain damaged, 15 left-brain damaged, and 16 control subjects. Emotional states were induced in subjects by having them view pleasant and unpleasant slides, for example, a beautiful sunset or an unpleasant surgical procedure. After each slide, subjects were videotaped while talking about their feelings and reactions to the slide. Independent judges evaluated the videotapes on a variety of criteria, results of which revealed that RBDs were more propositional than prosodic and more descriptive than affective relative to both NCs and LBDs suggesting to the authors that the right hemisphere plays a dominant role in emotional expression.

Gorelick and Ross (1987) examined spontaneous and posed prosody production as well as comprehension of prosody in 14 RBD patients at approximately two weeks

post-onset of stroke. Spontaneous speech was assessed during causal conversation and upon discussion of emotionally charged events in the patient's life. Observations were made about the affective-prosodic quality of the patient's voice and any mismatching of the affective components of language to the semantic content. Patient's speech was judged to have either normal affective prosody or moderate or severe loss of affective prosody. Patients were then classified into clinical categories based on this evaluation, their performance on the affective-prosodic repetition and comprehension tasks, and lesion site. Based on this categorization, Gorelick and Ross concluded that the right hemisphere was critical in aprosodia syndrome and its functioning parallels the left hemisphere's organization for language.

Ross et al. (1997) administered an aprosodia test battery to 10 LBDs and 12 RBDs over a 3-year period. Spontaneous affective prosody was assessed by interviewing subjects for 5 to 10 minutes and encouraging them to talk about affectively laden personal life events including a) reactions to being ill and hospitalized; b) the death of a close friend; c) close calls with death; and d) happy events such as marriages or special occasions. The most pronounced affect was acoustically analyzed by taking as many distinct phrases as needed to sample 10 seconds of affectively driven speech. Posed affective production was also tested in 3 conditions including word repetition (e.g., repeating the sentence "I am going to the movies" in a neutral, happy, sad, surprised, angry, or disinterested intonation), monosyllabic repetition (e.g., repeating the "sentence" "ba ba ba ba ba ba" in a neutral, happy, sad, surprised, angry, or disinterested intonation), and asyllabic repetition (e.g., creating different intonation using sentence: "aaaaaaahhhhhh"). Acoustical analysis of the spontaneous speech condition revealed that both the LBDs and RBDs produced a lower mean F_0 than normal controls and were not

significantly different from each other. Again, the authors make the point that reducing verbal-articulatory conditions robustly improves the performance of left but not right brain damaged patients, supporting the supposition that affective prosody is strongly lateralized to the right hemisphere.

Finally, Langer et al. (1998) videotaped 10 LBDs, 11 RBDs and 7 NCs while engaging in social interaction with their spouse and an interviewer. Segments of their interactions were independently coded by two observers for social competency and personality, while nonverbal communication measures including tone of voice, facial expressivity, volume of voice, frequency of smiling, frequency of laughter, and frequency of eye contact were self-rated and rated by the subjects spouse. Unfortunately, aggregate scores were analyzed so results regarding prosody were not specifically reported, but RBDs were found to be significantly less emotionally expressive compared to NCs, and LBDs were not significantly different from either RBDs or NCs. Other analyses suggested that LBDs were less socially competent than RBDs and NCs. In sum, the authors concluded that RBDs were less emotionally expressive compared to LBDs who demonstrated greater social impairments.

Based on this review, only six case studies (Blonder et al., 1995; Dykstra et al., 1995; Lebrun et al., 1985; Leentjens et al., 1998; Ross & Mesulam, 1979), one study evaluating RBDs alone (Gorelick & Ross, 1987), and two studies including RBDs, LBDs, and normal controls (Borod et al., 1985; Ross et al., 1997) have evaluated the neural mechanisms of spontaneous affective prosody. Besides the paucity of literature prohibiting meaningful assessment of the neuropsychological theories of spontaneous affective prosody production, the literature that is available is incomplete. Gorelick and Ross (1987) used subjective clinical evaluations to categorize right brain-damaged

patients which are unreliable (Baum & Pell, 1999) and did not include left brain-damaged patients or normal controls for comparison purposes. Borod et al. (1985) demonstrated prosody deficits in RBDs relative to other groups but did not quantify those differences. Ross et al. (1997) also evaluated spontaneous affective speech, but this was incomplete because the authors only evaluated fundamental frequency differences produced by RBDs and LDBs relative to normal controls. Several studies demonstrated the importance of acoustical cues in addition to fundamental frequency for the proper communication of specific emotions (e.g., (Pell & Baum, 1997; Van Lancker & Sidtis, 1992)). Further, it has been shown that human raters are able to detect emotive differences in speech when acoustical analyses fails to discriminate between emotions (Sobin & Alpert, 1999) suggesting that human ratings of spontaneous speech may reveal differences that might go undetected by acoustical analysis.

Acoustical profiles of specific emotions have not been completed for spontaneous affective speech. Ouellette et al. (Ouellette & Baum, 1994) performed an analysis of spontaneous linguistic prosody, but not affective prosody and Pell et al. (1999a; 1999b) evaluated posed output but only compared RBDs to matched controls. Excluding LDBs, this precluded any evaluation of the differential cue lateralization hypothesis, of which one version suggests that the left hemisphere controls duration cues with the right hemisphere controlling fundamental frequency cues. Studies that have examined the relationship between acoustical cues present in speech and human rating emotion misidentifications, also typically use posed, not spontaneous speech (Banse & Scherer, 1996; Gandour et al., 1995). They also fail to quantify specific parameter fluctuations based on rater judgments for particular emotions.

Current study

Spontaneous happy, sad, angry, and neutral speech samples were elicited from right brain-damaged, left brain-damaged, and matched control subjects and were judged by naïve raters for emotional intensity, identification, and identification confidence. Each speech sample also underwent acoustical analysis to measure pitch, duration, and amplitude parameters. Together, these data enabled the examination of emotional communication hypotheses regarding prosody including the right-hemisphere, valence, subcortical, and cue-lateralization hypotheses. Additionally, the relationship between prosodic cues present in speech and rater perception was examined.

The right hemisphere hypothesis predicts reduced intensity and impaired identification ratings of all emotional output by RBDs relative to the output of both LBDs and controls. The valence hypothesis predicts RBDs should have reduced intensity and/or impaired identification ratings for "sad" and "angry" output relative to LBDs and controls, whereas LBDs have reduced intensity and/or identification for "happy" output relative to RBDs and controls. Support for the existence of exclusive subcortical mechanisms for affective prosodic output would come from findings of reduced intensity and/or impaired identification ratings for patients with subcortical lesions, without evidence of deficits in subjects with cortical lesions to either hemisphere. The cue-lateralization hypothesis predicts that LBDs should show impairment using duration cues in prosody relative to RBDs and normal control subjects, and that RBDs use fundamental frequency cues in prosody relative to LBDs and normal control subjects measures (Baum, 1998; Baum, Pell, Leonard, & Gordon, 1997) (Baum & Pell, 1999).

Relationships between acoustics and accurate emotional identification of the speech samples were also explored. Banse and Scherer (1996) characterized happiness as

having a relatively average level of Fo and mean energy combined with increased speech rate relative to other emotions, while Sobin and Alpert (1999) found joy to be negatively associated with volume variance and positively correlated with pitch variation and duration of speech. This study also investigated whether increased ability to accurately identify happy, sad, and angry was related to increases or decreases of acoustic parameters.

Methods

Participants

Overview. Fifty five individuals served as subjects in this study and were a subset of brain-damaged and normal control subjects who had participated in a larger NIH-sponsored study (Joan C. Borod, Ph.D., Principal Investigator) at Mount Sinai Medical Center. Spontaneous emotional monologues were elicited from these individuals to assess their ability to express emotion in natural conversation. Additionally, 14 individuals served to rate numerous short spontaneous speech samples produced by the subjects, and will therefore be called "raters".

Subjects. Of the 55 subjects, 37 were individuals with unilateral brain damage (BDs), 17 of whom had damage to the left hemisphere (LBDs) and 20 to the right hemisphere (RBDs); 18 were neurologically normal and served as normal controls (NCs). Brain-damaged subjects had unilateral strokes as confirmed by CT and/or MRI. RBDs and LBDs were eligible for the study if they had a single-episode cerebrovascular accident. Prior to participation in the larger NIH study, verification of stroke, relevant neurological history, and lesion site were determined by the clinical medical record, the neurologist's report, and/or, the report of the neuroradiologist at the time of initial hospitalization.

In order to select subjects to be included in the study, data from all brain-damaged subjects were reviewed to ensure that the relevant monologues (i.e., participant's recount of the last time or a recent time they felt happy, sad, and angry as well as the last time or a recent time they went to the supermarket or bought something for the house) and control measures were completed and that the quality of the tapes was sufficient for evaluation (e.g., audio not clouded by static, all tapes available). Based on the anatomical and literature review of prosody (see page 19), all brain-damaged subjects with lesions to frontal (cortical-striatal PAG pathway), temporal (amygdala/PAG pathway, auditory cortex), and to test the relevant theories of emotion subcortical basal ganglia were included in the study. Subjects with frontal pathology were classified as "frontal" regardless of where additional lesions were located including temporal and subcortical structures. "Temporal" subjects also could have additional cortical or subcortical lesions, provided they were not frontal in nature. "Basal ganglia" subjects did not have additional cortical pathology but were permitted to have subcortical white matter lesions. Using this classification system, 7 LBDs and 11 RBDs with frontal pathology, 4 LBDs and 4 RBDs with temporal lesions, and 6 LBDs and 5 RBDs with subcortical basal ganglia lesions were included.

Table 1: Lesion location of subjects

Frontal +														
<i>Subject</i>	<i>Cortical Sites</i>				<i>Subcortical</i>									
	<i>F</i>	<i>T</i>	<i>P</i>	<i>O</i>	<i>nos</i>	<i>IC</i>	<i>Extrn</i>	<i>CR</i>	<i>Claus</i>	<i>Caud</i>	<i>Put</i>	<i>GP</i>	<i>BG</i>	<i>TH</i>
207	1													
208	1					1							1	
219	1													
220	1							1						
222	1		1					1						
227	1	1												
229	1		1										1	
234	1													
235	1		1			1	1	1	1	1	1			
240	1		1	1										1
299	1		1											
RBD Total	11	1	5	1	1	1	1	3	1	1	1		2	1
308	1	1	1				1							
311	1		1		1									
318	1		1											
331	1	1				1	1		1					
338	1							1						
339	1							1						
342	1					1					1			
LBD Total	7	2	3		1	2	2	2	1		1			
Temporal +														
<i>Subject</i>	<i>Cortical Sites</i>				<i>Subcortical</i>									
	<i>F</i>	<i>T</i>	<i>P</i>	<i>O</i>	<i>nos</i>	<i>IC</i>	<i>Extrn</i>	<i>CR</i>	<i>Claus</i>	<i>Caud</i>	<i>Put</i>	<i>GP</i>	<i>BG</i>	<i>TH</i>
201		1												
237		1		1										1
242		1	1											
246		1												
RBD Total	4	1	1											1
315		1		1			1							
329		1												
335		1												
341		1												
LBD Total	4		1				1							
Subcortical Gray Matter														
<i>Subject</i>	<i>Cortical Sites</i>				<i>Subcortical</i>									
	<i>F</i>	<i>T</i>	<i>P</i>	<i>O</i>	<i>nos</i>	<i>IC</i>	<i>Extrn</i>	<i>CR</i>	<i>Claus</i>	<i>Caud</i>	<i>Put</i>	<i>GP</i>	<i>BG</i>	<i>TH</i>
200											1			
202											1	1		
RBD Total											2	1		

300																			1	
343																			1	
345																			1	
LBD Total																			2	1

Subcortical Gray Matter / Subcortical White matter

<i>Subject</i>	<i>Cortical Sites</i>				<i>Subcortical</i>									
	<i>F</i>	<i>T</i>	<i>P</i>	<i>O</i>	<i>nos</i>	<i>IC</i>	<i>Extrn</i>	<i>CR</i>	<i>Claus</i>	<i>Caud</i>	<i>Put</i>	<i>GP</i>	<i>BG</i>	<i>TH</i>
225						1		1		1				
231						1		1			1	1	1	1
233						1				1				
RBD Total						3		2		2	1	1	1	1
317						1								1
320							1	1	1		1			
337						1							1	
LBD Total						2	1	1	1		1		1	1

A group of normal control subjects (n=18) were matched based on relevant demographic factors. A process was carried out in which a subset of the 94 NC subjects from the larger NIH study that had data for the spontaneous expression tasks were chosen. These subjects' data were checked for completeness and quality before being added to the final group from which the NCs would be drawn. Then, a group of several possible NCs was chosen based on the demographics that are considered relevant in studies of emotional expression, i.e., age, education, occupational status (as measured by the Hollingshead scale (Hollingshead, 1977), gender, and ethnicity (based on NIH definitions) and their overall demographics were calculated. Upon seeing the results of this process, the experimenter and the principle investigator of the larger study determined what changes were necessary to make the groups better matched (e.g., the NCs were initially too old and their education an average of two grades higher than BDs). NCs were then added and deleted one-by-one until the groups were very well matched. Finally, a group of 18 NCs was selected and demographics determined.

Table 2: Demographics by Gender

Group	Count	gender	AvgOfage	AvgOfHoll	AvgOfEDU	AvgOfMPO	StDevOfMPO
NC	11.00	Male	66.18	6.45	14.91		
NC	7.00	Female	65.14	5.86	14.00		
RBD	11.00	Male	69.64	6.73	14.45	64.73	73.62
RBD	9.00	Female	65.44	4.89	13.11	34.78	43.25
LBD	12.00	Male	65.25	5.67	14.42	11.92	11.11
LBD	5.00	Female	63.00	5.00	14.40	44.60	61.15

Table 3: Demographics

Group	Count	Average age	Average Holl	Average education	Average MPO	Median MPO
1	18.00	65.78	6.22	14.56		
2	20.00	67.75	5.90	13.85	51.25	29.50
3	17.00	64.59	5.47	14.41	21.53	9.00

All subjects completed an IRB-approved consent process. All were remunerated for their time and compensated for travel expenses. In order to compare demographic characteristics of the three subject groups, we conducted one way analyses of variance (ANOVAs) for Group (RBD, LBD, and NC) on each characteristic. No significant group differences were found for age, $F(2, 49) = .487, p = .617$; years of education, $F(2, 49) = .265, p = .769$; or occupational status, $F(2, 49) = 0.786, p = .461$. There was no significant difference in months post onset of stroke between the two brain-damaged subject groups $F(1,33) = 1.470, p = .234$. There were no significant differences in gender across the three groups, $\chi^2(1, 55) = 0.952, p = .621$, nor differences in ethnicity, $\chi^2(1, 55) = 0.06, p = .969$.

Table 4: Ethnicity

Ethnicity	Group	Count
White	NC	11
White	RBDs	13
White	LBDs	10
Non-White	NC	7
Non-White	RBDs	7
Non-White	LBDs	7

Table 5: Subject Demographics

SubID	age	Ethnicity	Hollingshead	Years education	MPO
102	56.00	1	6.00	18.00	
109	60.00	1	4.00	13.00	
114	80.00	1	6.00	14.00	
120	76.00	1	9.00	20.00	
121	71.00	1	7.00	18.00	
123	75.00	1	8.00	16.00	
127	71.00	1	6.00	15.00	
133	66.00	1	5.00	12.00	
136	75.00	1	6.00	15.00	
138	50.00	5	5.00	15.00	
157	60.00	5	6.00	13.00	
160	71.00	3	6.00	13.00	
183	56.00	1	6.00	11.00	
193	57.00	1	6.00	12.00	
410	75.00	2	6.00	13.00	
421	63.00	5	6.00	12.00	
426	52.00	5	6.00	14.00	
430	70.00	3	8.00	18.00	
200	70.00	1	9.00	17.00	180.00
201	50.00	1	7.00	16.00	31.00
202	78.00	1	6.00	11.00	8.00
207	72.00	1	9.00	20.00	27.00
208	62.00	1	4.00	14.00	225.00
219	65.00	1	9.00	18.00	41.00
220	75.00	5	5.00	14.00	3.00
222	50.00	5	3.00	12.00	3.00
225	60.00	1	7.00	16.00	35.00
227	78.00	5	7.00	11.00	65.00
229	71.00	1	6.00	12.00	9.00
231	72.00	1	4.00	12.00	28.00
233	75.00	5	6.00	18.00	43.00
234	76.00	5	1.00	5.00	11.00
235	52.00	1	4.00	12.00	9.00
237	72.00	1	5.00	12.00	9.00
240	69.00	1	6.00	12.00	60.00
242	89.00	5	8.00	14.00	12.00
246	63.00	5	5.00	15.00	144.00
299	56.00	1	7.00	16.00	82.00
300	62.00	1	5.00	16.00	149.00
308	71.00	1	7.00	18.00	10.00
311	69.00	5	4.00	12.00	30.00
315	64.00	1	4.00	8.00	2.00
317	62.00	1	9.00	19.00	4.00

SubID	age	Ethnicity	Hollingshead	Years education	MPO
318	73.00	5	6.00	13.00	6.00
320	54.00	6	2.00	12.00	16.00
329	70.00	5	1.00	12.00	39.00
331	52.00	1	5.00	13.00	49.00
335	48.00	5	5.00	14.00	6.00
337	68.00	1	6.00	12.00	5.00
338	79.00	1	8.00	16.00	12.00
339	75.00	1	4.00	11.00	9.00
341	74.00	1	7.00	18.00	3.00
342	53.00	5	5.00	16.00	8.00
343	50.00	5	6.00	15.00	11.00
345	74.00	1	9.00	20.00	7.00

All subjects were right-handed by self-report and according to the Coren, Porac, and Duncan (1979) lateral preference inventory, without a history of being forced to switch from left- to right-handedness. All subjects were native English-speakers or had begun using English by age 7. Prior to being entered into the study, all subjects had been subjected to careful phone screening to eliminate those with any history of prior or current neurological insult or disease, psychiatric disorder, mental retardation, substance abuse or learning disability. Upon beginning the study, all subjects were interviewed with the Schedule for Affective Disorders and Schizophrenia-Lifetime Version (SADS-L; Endicott & Spitzer, 1978) to exclude any participants with significant premorbid or current psychiatric disorders or any history of significant substance abuse.

All subjects met pre-established entry criteria based on cognitive screening tasks to ensure they possessed the basic cognitive ability to complete the experimental tasks of the New York Emotion Battery (NYEB; (Borod, Welkowitz, & Obler, 1992)). This battery was developed to assess emotional functioning using a componential approach. Subjects are asked to perform emotional tasks involving facial, lexical, and prosodic channels of emotion including happy, sad, anger, pleasant surprise, unpleasant surprise,

fear, interest, and disgust. Nonemotional control conditions are also included for each channel and mode (e.g., elicitation condition).

The Block Design and Information subtests of the Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981) were administered to assess basic intellectual skills necessary to adequately complete the experimental tasks. The Block Design subtest was administered to the NCs and the LBDs as a non-verbal test and the Information subtest was administered to the NCs and RBDs; these evaluated premorbid intellect in a form least likely to be hampered by the subject's stroke. Two subtests (Attention and Memory) of the Mattis Dementia Rating Scale (MDRS; Mattis, 1988) were administered to all subjects to assess basic memory and attention skills. One subtest (Complex Ideational Material) of the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) was administered to ensure that all subjects met the basic linguistic requirements (e.g., understanding instructions). The cutoff scores used for the screening measures in this study were generally 1-2 standard deviations (Sds) below the normal mean. The SADS-L served as the screening measure for mood and the Beck Depression Inventory was used to assess current mood in all subjects with no preset cutoff point.

Power Analysis Breitenstein et al. (1998) used scores from the Florida Affect Battery to compare 8 left anterior, 8 left posterior, 8 right anterior, and 8 right posterior subjects, to 10 healthy controls. This study was used as a basis for a power analysis because it compared brain-damaged subjects similar in lesion location to those used in the current proposal on a battery that included emotional prosody. Using the mean differences found in Breitenstein et al. (1998), a power analysis based on 17 subjects

revealed a likelihood of .997 ($DF_1 = 2$, $DF_2 = 50$, $F^* = 3.18$, $\lambda = 25.02$) to reveal differences between brain-damaged groups and normals, should differences exist.

Sobin and Alpert (1999) had 38 normal subjects produce happy, sad, angry, and fearful intonation while saying standardized sentences. An acoustical analysis was performed on the subjects' production, including measures proposed in the current proposal. Using the mean differences between the emotional conditions from Sobin and Alpert (1999) as a basis, a power analysis using 17 subjects revealed a likelihood of 1.00 for mean F_0 ($DF_1 = 2$, $DF_2 = 50$, $F^* = 3.18$, $\lambda = 206.89$), 0.45 for duration of speech ($DF_1 = 2$, $DF_2 = 50$, $F^* = 3.18$, $\lambda = 4.79$), and .998 for variance of volume ($DF_1 = 2$, $DF_2 = 50$, $F^* = 3.18$, $\lambda = 25.92$). Several other studies failed to provide an appropriate basis for a power analysis as they did not include means or variance information. These studies included Shapiro and Danly (1985), Cancelliere et al. (1990), Pell and Baum (1997), Pell (1999c), Van Lancker and Sidtis (1992), and Ross (1997). In sum, current power analyses indicated a high likelihood that statistical tests will be sensitive enough to detect differences between the proposed groups and conditions, should differences, in fact, exist.

Raters. Individuals who judged the spontaneous speech output are referred to as "raters". Raters were 14 undergraduate female students at Queens College who responded to posted advertisements and/or classroom presentations by the experimenter describing the responsibilities of a rater. All raters were paid the same hourly wage (\$8.50 per hour) for their time. Raters were subjected to screening procedures, similar to those used with the subjects, to ensure no history of neurological insult, significant substance abuse, major psychiatric disorder, or learning disability. All raters were native

English speakers or had begun to use English by age 7. The average age of raters was 20.4 years.

Collection of Expressive Data from Subjects

Experimental Tasks

The subjects produced spontaneous discourse during the monologue expression task from the New York Emotion Battery (NYEB; (Borod et al., 1992)). Subjects were informed that the purpose of the research was to study emotional communication and were aware that they were being video/audiotaped. The sampled emotions were derived from those described by Ekman and Friesen (1975) and Izard (1977), including happy, pleasant surprise, interest, sad, anger, unpleasant surprise, disgust, and fear. The scope of emotions evaluated in this study were limited to happy, sad, and angry in order to reduce the variability of emotion identification judgments as well as create a manageable dataset. Furthermore, these three emotions in particular permitted the evaluation of the 3 major emotional processing hypotheses (i.e., right hemisphere hypothesis [RH dominant for all 3 emotions], valence hypothesis [LH Dominant for Happy; RH dominant for Sad and Angry], and motoric direction hypothesis [LH dominant for Happy and Angry; RH dominant for Sad]).

Subjects were instructed to talk about emotional experiences involving each emotion, one at a time. They were to try to remember and actually relive the last time that they felt the target emotion, to put themselves back into the situation, and to try to re-experience it with as much real feeling and intensity as when it actually occurred. Subjects began to speak as soon as they could vividly remember and feel the experience and were instructed to speak for at least two minutes but not more than four minutes. At the end of each monologue, subjects were asked to report how intensely they re-

experienced the emotion as they told the story and how accurately they were able to convey the emotion. Subjects addressed each target emotion in 1 of 4 different orders to control for emotional carryover effects among subject samples.

If the subject had difficulty remembering the "last time", the experimenter prompted the subject with the cue: "If you can't remember the last time, you may tell me about the most recent time you can remember." Further, if the subject was having difficulty remembering any time they felt the target emotion, the examiner provided one, and only one, example from a list of examples for each emotion to help the subject remember an event from their own life that was applicable. For example, if a subject was having difficulty thinking of a time they felt the emotion of anger, the experimenter read an example from a list such as, "You were in a rush to get somewhere and a bus pulled out in front of your car", to the subject.

After hearing the example, the subject was instructed to again think back to the last time or a recent time that they had felt anger in their own life. Once the subjects began their monologue, the experimenter was quiet and did not interrupt the story or ask questions. However, if subjects were having difficulty speaking for at least 2 minutes, the experimenter cued the subject with neutral prompts, such as: "Go on.", "Can you tell me more about that?", "Anything else", "Try to describe what happened in greater detail".

Nonemotional Control Tasks Nonemotional spontaneous

monologues are elicited for control purposes as part of the New York Emotion Battery (Borod et al., 1992). Eight monologues are produced using relatively similar instructions as those for the emotional monologues, non emotional targets include: going for a walk, going to a restaurant, buying something for the house, making a meal, shopping at the supermarket, watching TV, visiting someone, and taking a form of transportation.

Zgaljardic (2000) asked ten trained battery administrators their impression of which monologue topics elicited the least emotional content when they gave the tests. Pilot data indicated that “shopping at the supermarket” ($M = 2.4$ (1 to 10 scale, 1 being less emotional); $SD = 1.07$) was the “least” emotional of the eight choices and was used in the current study to sample neutral spontaneous speech for baseline purposes. The second least emotional monologue was “buying something for the house” ($M = 2.4$ (1 to 10 scale, 1 being less emotional); $SD = 1.78$) and was used for the one subject that did not produce a “shopping at the supermarket” monologue.

If the subjects were having difficulty remembering what they did the last time they went grocery shopping, the experimenter gave them one cue to help them remember and tell their own experience. Example cues used for shopping at the supermarket include: a) “You were doing your regular grocery shopping. You got a shopping cart and started with the vegetable section and worked your way through each aisle until you got what you needed”; and b) “The aisles in the store were rearranged and the new wall covering gave it a much brighter appearance”.

Experimental Procedures

Overview The experimental procedures included several steps: 1) two experimenters selecting speech samples from the monologues; 2) digitizing subject monologue output; 3) creating presentable speech samples of 10- to 20-seconds in length; 4) low-pass filtering the speech sample to eliminate lexical information from the sample; 5) standardizing the volume of each sample to approximately 70 db; 6) presenting the speech samples to judges for ratings; and 7) conducting an acoustical analysis on the speech samples after which the hypotheses of the experiment were assessed through

statistical analysis. Most of these procedures were piloted, the results of which have been included in Appendix A for reference.

Monologue Sampling Our goal was to acquire a spontaneous speech sample that represented the subjects' ability to express the target emotion prosodically. One of the benefits of the monologue procedure is that it allows subjects to control the content of their discourse, providing a realistic spontaneous speech sample. However, due to the free nature of the monologue, subjects frequently include background to the story as well as other emotional events leading up to the target emotion event that would skew the speech sample if analyzed as part of the target emotion sample. For example for the target emotion "angry", one subject explained how his brother died and how that was sad (and it sounded sad) in order to explain why he had to sell a piece of property. His experience with the realtor made him angry, and he sounded angry as he went on to explain that experience. Therefore, it was critical for the validity of the study to have a mechanism to take a sample of, in this case, the angry discourse leaving out the sad and other extraneous discourse that would misrepresent the speech sample.

Another difficulty in acquiring an emotionally accurate speech sample is that there is not a preset time in a monologue that the target emotion will be expressed. Additionally, the possibility of invalidating the study by allowing the experimenters to evaluate the emotion of the monologues to create speech samples was of concern. As such, we developed sampling criteria utilizing written transcripts of the monologue. The story's contextual structure was thereby used as an unbiased means of selecting the portions of the monologue that contained speech samples of the target emotion, should a sample exist. All the discourse that was most likely to contain prosodic samples of the target emotion were included for analysis while the discourse that was most likely to

contain samples of neutral or non-target emotional discourse was eliminated from analyses. There are several advantages to this approach. The first is that the samples were not based on hearing the prosody or other non-verbal cues (e.g., facial, gestural, and/or body posture) used to express emotion during the monologue. The second is that it relieves the burden of having to pinpoint “the best” sample of emotion for each monologue.

Monologue sampling procedure.

Categories

The monologue sampling procedure categorized the discourse into four basic parts based on the story structure, which is not always sequentially ordered in the monologue. These were:

- a) Context: Information leading up to the emotion eliciting event or context to the event that furthers the understanding of the story.
- b) Play-by-play: The actual event or series of events that produced the target emotional experience.
- c) Reflection: Discourse that expressed what the experience meant to the subject, why it made them feel the way they did, or how the specific experience represented their feeling of the target emotion in an abstract sense.
- d) Closing: What happened after the event and/or summing up the story, independent of why the subject felt the way they did.

Discourse that conveyed the “play-by-play” of the event that produced the target emotion, and any “reflection” was included for analysis. These two types of segments were included as the most likely to include emotional content, should the subject convey emotion at all during the monologue. On the other hand, discourse that provided

“context” was excluded from rating. Further, “closing” was also excluded, provided the discourse was *not* related to how the subject felt or what the event meant to them, in which case it is “reflection” and was included. Several criteria have been created to help categorize the monologue into these four components and are explained in detail below (See Appendix B for full details).

Criteria for Categorization

Criterion Number 1: Temporal. All the events that were conveyed in the monologues are, by definition, from the past. However, since the subjects were asked to recall a specific point in time they felt a particular emotion, the story usually temporally revolves around the emotional event(s). Generally, discourse that should be categorized as “context” occurs temporally before the event that produced the target emotion. The “play-by-play” occurs in the present tense of the story and the “closing” occurs in the future or includes statements unrelated to the target event. “Reflection” statements are generally not time specific, and it is therefore crucial to identify the target event to be able to assess whether the reflective comments refer to the target event or to something else.

Criterion Number 2: Identifying the Key Event. The key event of the monologue is the situation or activity that the subject reports produced the target emotion. Identifying the key event helps categorize the discourse correctly, particularly for the “reflection” categorization. In cases where the key event is ambiguous, the temporal criterion takes priority. Refer to Appendix B for specific examples of how the Temporal criterion facilitated the identification of the Key Event criterion in some cases.

Criterion number 3: Experimenter cues. The last criterion relates to discourse following cues from the experimenter. Although the experimenters had clear guidelines

regarding giving cues, occasionally these guidelines were not strictly followed.

Discourse following inappropriate cues by the examiner (see above) was eliminated from discourse analysis.

Conferencing.

The two experimenters resolved any discrepancies through conferencing which included the following steps: 1) Identify a segment that is not agreed on and discuss each experimenter's categorization of the discourse; 2) Review the entire monologue to determine if both experimenters are "interpreting" the monologue similarly (e.g., main topic and what was "present" vs. "background" of the story). If one interpretation was agreed to be better than another, the segments in question were re-marked based on this new interpretation); 3) Discrepancies that still existed were reviewed in detail in relation to the temporal and primary event criteria, the former taking precedence, and a new category decision was made about the segment based on what was agreed to fit the monologue best. All discrepancies were resolved with the above methods. Had they not been, a pre-identified third person would have been brought in for consensus.

Digital processing of speech samples.

Selected speech samples were digitized from audio/video tape to computer through a sound blaster soundcard at 22,000 Hz. As a rule, samples were cut based on the sentence structure into 10- to 20-second audio samples using Creative Wave Studio 4.08, such that no sentence was cut into parts. When it was necessary, some samples were slightly less than 10 seconds or slightly more than 20 seconds in order to maintain the integrity of a sentence or sentences within a single sample. Only 6.2% of the 796 samples were less than 9 seconds and 5.9% were greater than 20 seconds.

Speech samples were low-pass filtered to eliminate 90-95% of the lexical content of the speech, leaving the prosodic intonations essentially intact (Speilman, 2002): this prevented verbal material of the sample from influencing the perception of prosody. Typically, male and female voices are filtered at different levels (e.g., males filtered at 400Hz, females filtered at 600Hz) due to the well-established fact that male mean frequency levels are significantly lower than females (Baken & Orlikoff, 2000). However, filtering samples at two standard rates fails to address the intra-sex, and inter-emotion fundamental frequency differences effectively. Therefore, Praat 6.08 speech analysis software was used to calculate the average fundamental frequency of each individual's samples and then the samples were low-pass filtered based on the conversion rates found in Table 6.

Table 6: Voice fundamental frequency/low-pass filter level conversion table.

Voice Fo	Filter Hz	Voice Fo	Filter Hz	Voice Fo	Filter Hz	Voice Fo	Filter Hz	Voice Fo	Filter Hz
80	300.0	120	400.0	160	500.0	200	600.0	240	700.0
85	312.5	125	412.5	165	512.5	205	612.5	245	712.5
90	325.0	130	425.0	170	525.0	210	625.0	250	725.0
95	337.5	135	437.5	175	537.5	215	637.5	255	737.5
100	350.0	140	450.0	180	550.0	220	650.0	260	750.0
105	362.5	145	462.5	185	562.5	225	662.5	265	762.5
110	375.0	150	475.0	190	575.0	230	675.0	270	775.0
115	387.5	155	487.5	195	587.5	235	687.5	275	787.5

This table was created by pairing the average fundamental frequency for male (e.g., 120 Hz) and female (e.g., 200 Hz) speakers (Baken & Orlikoff, 2000) with a typical filtering rate for male (e.g., 400 Hz) and female (e.g., 600 Hz) voices. This pairing can be seen in bold in the table. The rest of the table was extrapolated from these two core pairings, such that at a rate of every 5 Hz in voice equates to a 12.5 Hz adjustment in filtering. Using this flexible system enabled each sample to be filtered at an "ideal" level.

balancing the need to eliminate 90-95% of the speech content with leaving as much prosody intact as possible.

Finally, each sample was equalized for loudness using a script run in Matlab 6.01 which attenuates and/or amplifies each signal to reach an average of 70 dB. This was necessary because loudness affects perceived intensity and the samples varied widely in loudness, in part, due to testing conditions. This equalization removed loudness differences as a source of variation.

Rating Procedures

Measurement Introduction In order to evaluate the expressions produced by the three groups of subjects (NCs, RBDs, and LBDs), three ratings were taken: emotion identification, confidence in ones' identification rating, and the intensity of the emotion expressed. Emotion identification required raters to determine what emotion best fit the sample (i.e., angry, happy, sad). From these ratings, subjects' accuracy in expressing each emotion prosodically was inferred. To facilitate the interpretation of the emotional identification judgments, raters also indicated their level of confidence in the correctness of emotion label they identified for each sample. Finally, raters judged how intensely the emotion was expressed in the sample.

Rating Scales A 5-point likert scale (Borod et al., submitted for publication; Canino, 2001) was used to measure rater confidence and a 7-point likert scale (Canino, Borod, Madigan, Tabert, & Schmidt, 1999) was used to measure emotional intensity. Please refer to Table 7 for details on these scales.

Rater Aggregation: One method to obtain these judgments is to train a small number of highly reliable raters to identify specific aspects of the speech sample accurately. Training would consist of teaching raters to pay attention to specific features

of the sample, like speech rate. However, using specific criteria to make judgments on this dataset introduces several problems. First, it supposes that the experimenters can accurately define a universal definition of what each emotion should sound like. Second, raters trained to evaluate speech samples based on particular acoustical features renders analysis of the relationship between the rating data and the acoustical data meaningless since the raters have already been trained to listen to specific cues.

Therefore, a second method, aggregation, is much more appropriate for this dataset. Rather than training raters, raters are asked to make judgments based on their own criteria (i.e., without training). Aggregation relies on the inherent properties of averages and sampling to produce reliable data that can be generalized (Epstein, 1980). For judgments regarding emotion, it has been shown that 6 observations aggregated and compared to 6 observations, similarly aggregated, produces a stable behavioral measurement with a predicted reliability coefficient between .65 and .75 (Epstein, 1980). For this study, 2 groups of 7 raters (14 raters together) evaluated the speech samples. An average judgment value was created for each group, and a correlation was calculated between the two aggregated sample scores to assess overall agreement.

Table 7: Confidence and Intensity scales

Confidence							
	0	1	2	3	4		
	Not at all Confident	Somewhat Confident	Moderately Confident	Very Confident	Extremely Confident		
Intensity							
	0	1	2	3	4	5	
	Not at all Intense	Mimimally Intense	Mildly Intense	Moderately Intense	Quite Intense	Very Intense	Extremely Intense

Rating Session Procedure Individually, or in a small group of 2 or 3, raters were read a standard set of instructions detailing their task in the experiment and use of the computer program designed to present speech samples and record their responses (see

Appendix D for standard instructions). The program used by the raters to perform the experiment also contained a reference of the 3 scales (Identification, Confidence, Intensity) as well as a review of the major strategy points covered in the instructions. This information was available to the rater at any time by a simple button press and did not require her to stop the experiment or ask permission.

Before raters were able to perform the experimental ratings, a practice session was completed. The primary goal of the practice session was to allow self-training of what to listen for in the samples and to facilitate rater reliability by increasing intra-individual agreement on the emotional identification rating. Based on pilot work, seven speech samples were selected for use in the practice session. In the practice session, raters listened to each sample at random and recorded the emotion they thought each sample sounded like most. Afterwards, raters re-listened to the same seven samples in a new random order. Their responses from the first session and second session were compared. If raters agreed with themselves on 6 of the 7 samples, they were permitted to start the experimental session. If less than 6, raters listened to the same 7 samples in a new random order and their responses were compared to the previous set. This process continued until raters were in agreement with themselves at criterion levels. The practice session was performed at the beginning of every rating session unless a rater performed two sessions on a single day.

For the experimental session, there were 796 speech samples presented in blocks of 55 randomly selected trials. Typically, raters completed 1 to 3 blocks per rating session. Raters would start by pressing a button to listen to a sample. They could listen to each sample up to 3 times before being forced to make selections. The presentation program controlled the order of response by the rater and provided step-by-step

instructions that changed as the rater progressed through their selections. Raters first selected the emotion they thought the sample sounded like most. Once selected, the “confidence rating” area became active and the raters were asked how confident they were that the sample sounded like the emotion they selected. Finally, the “intensity rating” area became active and raters selected how intensely they thought the emotion they selected was expressed in the sample. Raters then had the opportunity to review their answers. If they changed a response, all previous responses were erased. For example, if raters went back and changed their confidence rating, their intensity rating would be erased and they would be re-asked the question. If they changed their identification rating, then both the confidence and intensity rating would be erased. Once the rater was satisfied with all three ratings, she would press a button to go on to the next trial to begin the process again. The presentation software recorded each rating and the number of times they listened to the sample for later processing.

Acoustical Analysis

Acoustical parameters. Praat 6.23 acoustical analysis software was used to extract acoustical measurements from all 796 speech samples. The autocorrelation method (Boersma, 1993) was used to extract pitch parameters including Fo mean (Hz), Fo standard deviation (semitones), and Fo acceleration. Additionally, the pitch value for every .01 second of the speech sample was exported and the pitch regression slope was calculated for the first half of the sample, the second half of the sample, and the entire sample (for review of the relevance of this issue, see (Pihan, 2001)). Several speech rate measures were calculated for each speech sample including the average number of syllables spoken per minute, average duration of voicelessness, the number of pauses less than 0.2 seconds, and the number of pauses greater than 0.2 seconds. Average syllables

uttered per minute was calculated using the written transcripts and was divided by the amount of voicing time (Banse & Scherer, 1996; Scherer, Banse, Wallbott, & Goldbeck, 1991; Sobin & Alpert, 1999). Finally, since mean speech volume has been digitally controlled, amplitude is not interpretable, and therefore, only amplitude standard deviation was measured.

Transformations To meet the assumption of linearity of errors for analysis, the acoustical measurements of pitch mean, pitch standard deviation, pitch slope, and volume standard deviation were log transformed.

Data Preparation

Monologue Sampling Procedure As described, this experimenter and another person from Dr. Borod's lab categorized all words using the monologue sampling procedure to select speech samples. Agreement on sample selection was calculated for each monologue by taking the total number of words agreed on and dividing it by the number of words in the monologue. Overall, three types of disagreement arose: 1) a word categorized "play-by-play" by one experimenter and "reflection" by the other; 2) a word being categorized as "not rated" (i.e., context, closing, or excluded due to cue) by one experimenter and "play-by-play" by the other rater; and 3) the same sentences being categorized as "not rated" by one experimenter and "reflection" by the other.

Disagreement rates were calculated for each type by taking the total number of words under each type and dividing by the total number of words in the monologue. The agreement and disagreement percentages for each sample were then used to generate an overall agreement rate for all monologues as well as an agreement breakdown by target emotion and the subject group.

Ratings Fourteen raters generated an identification, confidence, and intensity rating for all 796 speech samples. In order to assess questions regarding the emotion conveyed in the speech samples, two variables were created from these ratings. "Category" (levels: unknown, happy, sad, angry) and "Strength" (levels: 1, 2, 3) were calculated based on the number of raters that rated a speech sample as happy, sad, or angry. For the "category" variable, a sample was labeled with one emotion if at least seven raters selected that emotion and the other two emotion choices were endorsed by at least 3 fewer raters. If these two conditions were not met, the sample was labeled "unknown".

Each sample was also assigned a "Strength" level based on the difference between the two most selected emotion categories by raters for the sample. For example, if nine raters selected happy, four angry, and one sad, the difference between the two most selected categories is five (i.e., happy (9) - angry (4) = 5 rater difference). A "strength" value of "1" was given to a sample if the difference in rater count was between 3 and 5, "2" for a difference of 6 to 9, and "3" for a difference of 10 to 14. For samples labeled unknown, a difference of 0 was given a strength of "1", a difference of 1 was a strength of "2" and finally a difference of 2 received a strength of "3". For example see table 8:

Table 8: Examples of "Category" and "Strength" variable creation

Sample	# raters picking Happy	# raters picking Sad	# raters pick Angry	Assigned Category	Strength
1	6	0	8	Unknown	3
2	6	6	2	Unknown	1
3	3	9	2	Sad	2
4	1	1	12	Angry	3
5	8	4	2	Happy	1

For each sample, an average intensity and confidence scores was calculated for happy, sad, and angry based on the data produced by raters who selected the

corresponding emotion. This was necessary because the intensity and confidence ratings were dependent on the emotional category selected, so calculating a single intensity and confidence score for each sample was inappropriate. For example, if for a given sample 10 raters selected happy, 3 judges selected angry, and one selected sad, the average happy intensity and confidence ratings would be based on the ratings of 10 raters, whereas the averages for angry would be based on 3 raters, and the averages for sad would be based on 1 rater. A null intensity and confidence value was generated if none of the raters rated a sample a particular emotion.

In order to assess the subjects ability to convey prosodic emotion accurately, the "best" sample for each monologue was identified. The "best" sample was identified by the one sample from a given monologue for which the largest number of raters selected the target emotion. For example, if five samples were taken from a subject's Happy monologue, the "best" sample was the sample that had the most raters identifying it as 'happy' (i.e., the target emotion). In the event of a tie, the most intense of those samples was designated the "best" sample. For the control condition, the ratings from all the neutral samples were averaged together to establish a baseline level of intensity as they came from only one "target".

Data Analyses

Introduction Gricean pragmatics approaches discourse analysis from the perspective that successful communication of a spoken message is linked to the listener's interpretation of the speaker's intention (Bloom et al., 1999). As such, one way to analyze this data is from the perspective of the subject, with an emphasis on the emotion that the subject intended to convey. Another way to analyze this data is from the perspective of the rater emphasizing the raters' perception of the emotion in a sample.

Analyses based on the subjects' perspective is useful in answering questions regarding emotional communication hypotheses and to explore the how subjects' intend to convey emotion. Conversely, analyses based on the raters' perspective are useful to assess what speech sounds like when it successfully conveys emotion.

Monologue Sampling Reliability Descriptive statistics of experimenter agreement and disagreement rates were compiled. Additionally, a mixed-effects analysis was completed to assess whether initial experimenter agreement regarding samples differed based on subject group or monologue emotion. See the analysis description for "Acoustic-to-rating relationships" for full explanation of mixed-effects analysis.

Self-training Trial Data Descriptive statistics were calculated for initial self-training and overall self-training of the raters.

Rating Reliability Identification and intensity measurement reliability was calculated by dividing the 14 raters into two random groups of seven, averaging the rater data separately for the two groups, and measuring the Pearson r correlation coefficient between the two groups. This process was replicated 100 times with the assistance of a computer program that generated random combinations of rater groups. Therefore, the reported correlation coefficients are an average of 100 correlations calculated from these random combinations of rater groups.

Demographic influence on intensity and confidence ratings Multiple regression was used to assess whether subject demographic variables accounted for significant variation in rating measures. All demographic variables were entered into analysis simultaneously.

Subject Perspective - Rating Data To assess support for the right hemisphere, valence, and motoric direction hypotheses, analyses on identification, confidence, and

intensity ratings were conducted based on the subjects' "best" samples. The "best" sample produced for each monologue does not imply that most raters believed the sample actually sounded like the emotion the subject intended to convey. The sample is only the "best" available sample from that monologue providing minimal information regarding how accurately the sample represents the intended emotion. Analyses based on the "best" samples are useful to assess differences among the patient or gender groups in how they intend to express emotions, but are not necessarily an accurate indicator of how successfully emotions were conveyed.

Two types of Group x Gender x Emotion repeated-measures ANOVAs were conducted on each type of rating data. The first ANOVA was conducted using the neutral condition as a level of Emotion (Happy, Sad, Angry, Neutral). This usage of the neutral condition in analyses is hereafter referred to as a "neutral-level" analysis. This analysis is useful to establish differences in level between groups of subjects. The second used the neutral condition to create a difference score for each emotional condition (Emotion – Neutral, Sad – Neutral, Angry – Neutral). Using the neutral condition in analyses in this manner is hereafter referred to as a "difference-score" analysis. This analysis is useful to detect differences in emotional expression between groups. Together the two analyses paint a more complete picture than either analysis alone. Effects of gender and emotion were also evaluated on an exploratory basis and supporting analyses were also conducted.

Separate analyses for each rating data type were more appropriate than a doubly multivariate analysis because the correlations between the intensity and confidence ratings were quite high and the correlation between the identification ratings to the other two measures varied depending on the emotional condition. Multivariate analysis is most

appropriate when moderate correlations exist between the measures of interest and is less powerful than separate ANOVAs when this is not the case (Tabachnick & Fidell, 1996).

Subject Perspective - Acoustic Data Both “neutral-level” and “difference-score” Group x Gender x Emotion repeated-measures ANOVAs were performed on each acoustical parameter of the “best” samples to determine if there were acoustical differences in the samples that could support the findings of the rating data and/or potentially reveal differences not detected by human raters. These analyses also allowed direct testing of the cue-lateralization hypothesis (Left-hemisphere controls duration cues; Right-hemisphere controls pitch cues) in addition to traditional emotional communication hypotheses.

Rater Perspective - Acoustic to Category Relationships Fixed effects are parameters associated with an entire population and random effects are associated with individual experimental units drawn at random from the population. Fixed effects influence means and are used to predict population values, while random effects influence the variability of an effect (Pinheiro & Bates, 2000). In the current study, 796 samples were created from data of 55 subjects telling four stories. Samples taken from the same story or subject is a source of pseudoreplication and are not independent from one another. One way to deal with pseudoreplication is to average away replications; another is to perform a mixed-effects analysis.

Linear mixed-effects models were fitted to address the relationship between sample acoustics and the three rater judgments. This type of analysis allowed all 796 samples to be used and mixed-effects models handle missing observations and unbalanced data sets (as is the case here) more efficiently than classic repeated measures ANOVA (Bagiella, Sloan, & Heitjan, 2000). The linear mixed-effects (lme) model

available in S-Plus was used for these analyses because it is particularly flexible, but similar procedures are available through SAS PROC MIXED and in MIXED available in the newest version of SPSS 11.5.

To assess the relationship between sample acoustics and the rater-judged category of the sample, separate lme models using the restricted maximum likelihood (REML) method were performed using each acoustic parameter as the response variable. Fixed effects were Category (Unknown, Happy, Sad, Angry) x Strength (1,2,3) x Gender and the random effects were Subject x Monologue(Subject) (i.e., monologue nested within subject) to account for the pseudoreplication described above. The "varIdent" variance function was used to correct heteroscedasticity by gender and the "varCompsym" correlation structure was used to model the within-group dependence.

To assess the relationship between sample acoustics and the intensity and confidence ratings, separate lme models using the restricted maximum likelihood (REML) method were performed using the happy, sad, and angry intensity and confidence measures as response variables. The random effects in the analyses were Subject x Monologue(Subject) and the "varIdent" variance function was used to correct heteroscedasticity by gender. The "varCompsym" correlation structure was used to model the within-group dependence. As per Hosmer and Lemeshow (2000) and Pinheiro and Bates (2000), acoustical parameters were added as predictor variables one at a time until a significant predictor was found. Additional predictor variables were added to the model one at a time and were accepted only if ANOVA revealed that the Goodness-of-fit measures (Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)) were significantly lower, reflecting a better fit. Once an acceptable model was

found, discarded parameters were re-entered into the model to ensure that they should be excluded from the model.

Results

Monologue Sampling Reliability

Results show that experimenters LC and MS had an average agreement rate of 86.9% ($s = 13.9\%$) on 220 monologues, leaving 13.1% of the monologues in disagreement. Breaking down overall disagreement, 2.47% of the disagreement was attributable to segments being categorized as “play-by-play” by one experimenter and “reflection” by the other. The remaining disagreement was attributable to segments being categorized as “not rated” (i.e., context, closing, or excluded due to cue) by one experimenter and “play-by-play” by the other rater (8.69%) or categorized as “not rated” by one experimenter and “reflection” by the other (1.85%).

A Group (NC, RBD, LBD) x Emotion (Neutral, Happy, Sad, Angry) linear mixed-effects analysis revealed a main effect of Group ($F_{(2,52.85)}=4.062, p=.023$) and Group by Emotion interaction effect ($F_{(6,156.6)}=3.165, p<.006$). Post hoc comparisons revealed that the agreement rate on monologues of normal controls ($x = 82.87\%, s = 16.55\%$) was significantly lower than both RBDs ($x = 89.28\%, s = 11.83\%$) and LBDs ($x = 88.66\%, s = 12.50\%$). Additionally, sad monologues from normal controls were significantly less agreed upon by the experimenters than any other type (see table 9).

Practice Trial Data

In order to assess the effectiveness of the identification self-training, descriptive data was compiled from all the raters. For the initial self-training session, raters required an average of 3.07 additional blocks ($sd = 1.63$) before reaching the criterion level of 85% self-agreement rate (i.e., 6 out of 7 samples with identical identification ratings) to

Table 9: Emotion by Group monologue estimated mean agreement rates.

EMOTION	GROUP	Mean	Std. Error	df	95% Confidence Interval	
					Lower Bound	Upper Bound
anger	1	.811	.031	202.978	.749	.873
	2	.859	.030	203.335	.801	.918
	3	.914	.032	202.978	.850	.978
Happy	1	.909	.031	202.978	.847	.971
	2	.894	.030	203.335	.835	.952
	3	.887	.032	202.978	.823	.951
sadness	1	.728	.031	202.978	.666	.791
	2	.920	.030	203.335	.861	.978
	3	.882	.032	202.978	.818	.946
super	1	.866	.031	202.978	.804	.928
	2	.900	.030	203.794	.841	.959
	3	.864	.032	202.978	.800	.928

start an experimental session. The minimum number of additional self-training blocks was one and the maximum was seven, with two and four additional blocks representing the 25th and 75th percentiles respectively. Only one rater required more than four additional blocks to complete the initial self-training. Self-training was completed prior to every experimental session separated by a single day. The average number of self-training blocks required to reach criterion overall was 2.05 with a standard deviation of 1.92.

Rating Reliability

Reliability for the identification rating was broken down by emotion since it is a categorical variable. The reliability coefficient gives the average correlation of the number of raters selecting a specific emotion for each sample. Analysis revealed that the identification ratings were reliable, specifically, showing that raters were most reliable in identifying sad ($r = .7587$), followed by anger ($r = .6779$) and happy ($r = .6039$).

Breitenstein et al. (2001) found that recognition rates of emotion is generally best of (hot)

anger and sadness and worse for happiness suggesting that the reliability rates from this study are consistent with those found in the literature.

The same procedures were followed for calculating the reliability of the intensity ratings, except that an overall reliability, independent of emotion, was calculated and the emotional category of a sample was determined by which subject monologue it came from rather than what category the raters judged it to belong to as is the case in the identification rating. Intensity judgments for samples coming from happy monologues were the most reliable ($r = .6607$), followed by angry ($r = .6062$), supermarket [neutral] ($r = .5736$), and sad ($r = .5023$). Overall, the reliability of the intensity measurement was good ($r = .5922$).

Demographic influence on intensity and confidence ratings

To assess whether important subject demographic variables systematically varied with rater judgments, separate multiple regression analyses were conducted with emotional intensity and confidence ratings as the response variables. Predictor variables were entered into the analysis at once; these were subject education, socioeconomic status (SES), and months-post-onset of stroke (MPO). As shown in tables 10 and 11, MPO, subjects' education, and SES were not significant predictors of either rater intensity or confidence ratings and therefore were not included as covariates in subsequent analyses.

Subjects' Perspective - Rating Data

Correlations among Identification, Confidence, and Intensity Ratings Separate Pearson r correlations were calculated to assess the linear relationships between the identification, confidence, and intensity ratings for each emotional condition for the "best" sample data. For the happy condition, results revealed that as the number of raters

agreeing that as a sample was happy increased, rater confidence ($r = .593, p < .000$) and intensity ratings ($r = .467, p < .000$) increased. Additionally, there was a strong positive linear relationship between the intensity and confidence rating ($r = .847, p < .000$). For sad samples, the number of raters agreeing that a sample was sad was not significantly related to rater confidence ($r = .167, p = .223$) nor intensity ratings ($r = .195, p < .153$). However, as with happy, there was a strong positive linear relationship between the intensity and confidence ratings for sad samples ($r = .871, p < .000$). Finally, for angry samples, as the number of raters agreeing that a sample sounded angry increased, rater confidence ($r = .366, p = .006$) and intensity rating ($r = .399, p = .003$) increased.

Table 10: Multiple Regression – Demographic Prediction of Intensity Ratings

<u>Intensity</u>						
Happy		Model F = 0.34		Sig. = 0.796		
	<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>	
(Constant)	-.024	.513		-.048	.962	
education	3.376E-02	.053	.166	.634	.531	
HOLL	-.060	.081	-.193	-.746	.461	
MPO	1.215E-03	.002	.097	.558	.580	
Sad		Model F = 1.098		Sig. = 0.364		
	<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>	
(Constant)	.687	.375		1.834	.076	
education	-.046	.039	-.297	-1.172	.249	
HOLL	1.189E-02	.059	.050	.201	.842	
MPO	1.873E-03	.002	.199	1.179	.247	
Angry		Model F = 0.922		Sig. = 0.441		
	<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>	
(Constant)	.125	.448		.278	.783	
education	4.539E-02	.047	.249	.975	.337	
HOLL	-.113	.071	-.403	-1.598	.120	
MPO	3.484E-04	.002	.031	.183	.856	

Again, there was a strong positive linear relationship between the intensity and confidence ratings ($r = .860, p < .000$). Together, these results suggest that the relationship between the identification ratings to the other two ratings is dependent on the emotional

category. Further, the strong relationship between intensity and confidence suggest that separate univariate repeated-measures analyses are more appropriate than a single doubly multivariate analysis for the rating data (See Tables 10 and 11).

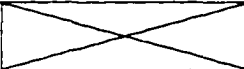
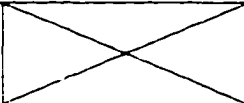
Table 11: Multiple Regression – Demographic Prediction of Confidence Ratings

Confidence						
Happy		Model $F = 0.349$		Sig. = 0.790		
		<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>
	(Constant)	-.016	.386		-.042	.966
	education	3.529E-02	.040	.230	.880	.385
	HOLL	-.056	.061	-.239	-.922	.363
	MPO	-.001	.002	-.084	-.482	.633
Sad		Model $F = 0.646$		Sig. = 0.591		
		<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>
	(Constant)	.459	.267		1.719	.095
	education	-.034	.028	-.321	-1.242	.223
	HOLL	2.593E-02	.042	.157	.616	.542
	MPO	7.431E-04	.001	.113	.656	.516
Angry		Model $F = 0.484$		Sig. = 0.696		
		<u>B</u>	<u>Std. Error</u>	<u>Beta</u>	<u>t</u>	<u>Sig.</u>
	(Constant)	.241	.355		.681	.501
	education	3.007E-03	.037	.021	.082	.935
	HOLL	-.047	.056	-.218	-.846	.404
	MPO	3.549E-04	.002	.041	.236	.815

Identification Ratings In order to assess support for the emotional communication hypotheses, a Group x Gender x Emotion (happy, sad, angry) repeated measures ANOVA was conducted using identification scores from subjects' "best" sample of happy, sad, and angry, independent of ratings from the neutral condition. A significant main effect of emotion ($F_{(2,98)}=8.461$, $p<.000$) and a significant Emotion x Gender interaction ($F_{(2,98)}=4.414$, $p=.015$) were revealed. Significant post hoc comparisons revealed that rater agreement was significantly higher for sad samples than for happy and angry. Further, rater agreements for happy and sad were higher for female subjects but

higher for males on angry samples. These results do not provide support for any of the emotional communication hypotheses (See Table 12).

Table 12: Rating data descriptive statistics broken down by Gender and Group

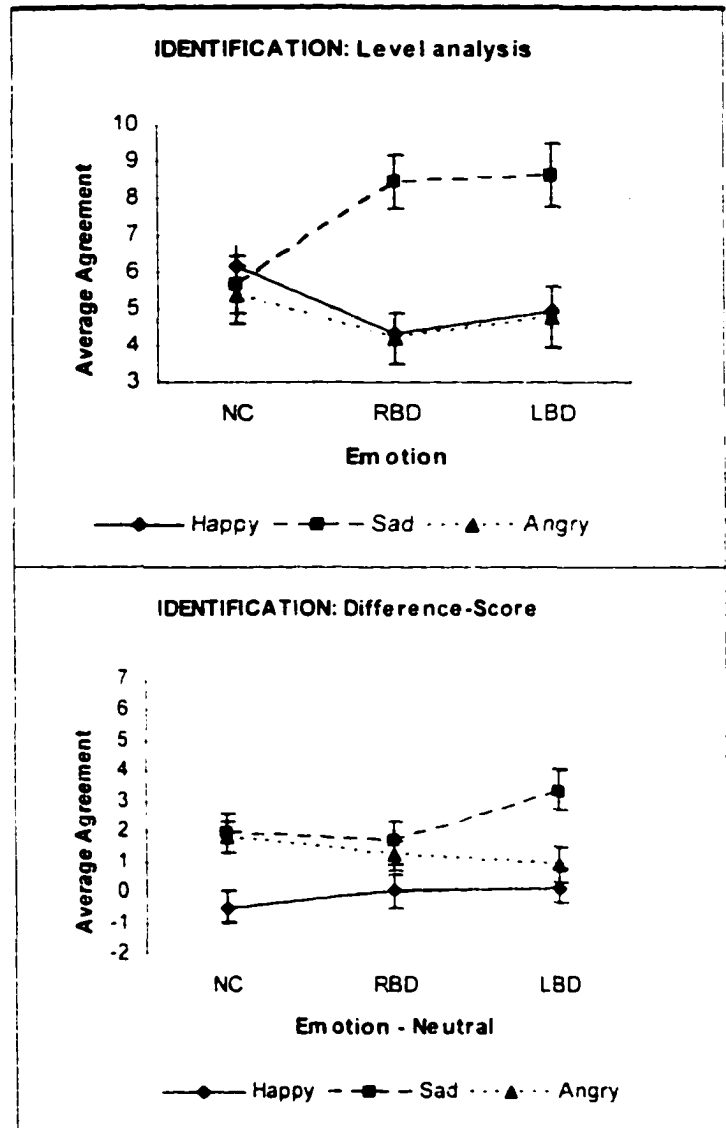
		Happy Mean \pm (SD)	Sad Mean \pm (SD)	Angry Mean \pm (SD)	Neutral Mean \pm (SD)
Identification	Men	6.15 \pm (2.79)	9.00 \pm (3.57)	7.74 \pm (3.50)	
	Women	8.48 \pm (3.09)	10.48 \pm (3.54)	6.10 \pm (4.25)	
Confidence	Men	2.83 \pm (0.40)	2.86 \pm (0.32)	2.92 \pm (0.37)	2.83 \pm (0.33)
	Women	3.34 \pm (0.55)	3.20 \pm (0.40)	3.08 \pm (0.54)	2.97 \pm (0.32)
Intensity	Men	3.68 \pm (0.55)	3.63 \pm (0.40)	3.78 \pm (0.45)	3.56 \pm (0.46)
	Women	4.22 \pm (0.68)	4.07 \pm (0.59)	4.06 \pm (0.77)	3.79 \pm (0.41)
Identification	NC	7.67 \pm (3.18)	8.28 \pm (3.43)	8.39 \pm (3.65)	
	RBD	6.35 \pm (2.48)	9.90 \pm (3.91)	6.10 \pm (3.85)	
	LBD	7.18 \pm (3.64)	10.53 \pm (3.18)	6.94 \pm (3.90)	
Confidence	NC	2.95 \pm (0.59)	2.83 \pm (0.42)	3.03 \pm (0.42)	2.80 \pm (0.38)
	RBD	3.01 \pm (0.45)	3.09 \pm (0.40)	3.02 \pm (0.48)	2.92 \pm (0.26)
	LBD	3.11 \pm (0.54)	3.05 \pm (0.29)	2.87 \pm (0.43)	2.93 \pm (0.35)
Intensity	NC	3.93 \pm (0.71)	3.60 \pm (0.42)	3.98 \pm (0.48)	3.51 \pm (0.48)
	RBD	3.82 \pm (0.61)	3.90 \pm (0.65)	3.87 \pm (0.63)	3.69 \pm (0.34)
	LBD	3.92 \pm (0.68)	3.89 \pm (0.41)	3.82 \pm (0.70)	3.74 \pm (0.50)

Although the Group \times Emotion interaction effect was not significant ($F_{(2,98)}=1.651, p=.167$), inspection of the data revealed that expressions of brain-damaged subjects resulted in higher than normal agreement rates for sad samples, but lower than normal agreement rates for happy and angry. Because of the clinical importance of this trend, additional analyses were conducted to explore whether it represents an actual effect or is spurious.

A Group \times Gender \times Emotion (happy, sad, angry) analysis was conducted using average scores from all samples because it was conceivable that brain damaged subjects' "best" sample might be statistically indistinguishable from normal controls' "best" sample, but significant differences could emerge when averages from all speech samples were considered. Results of the Group \times Emotion interaction were significant ($F_{(4,98)}=2.986, p=.023$) and significant post hoc comparisons revealed that rater agreement

was significantly higher for sad samples than for happy and angry. Additionally, a Group x Gender x Emotion (happy – neutral, sad – neutral, angry – neutral) analysis was conducted to assess the possibility that this effect represented an ability difference in emotional expression. The Group x Emotion interaction effect was not significant ($F_{(4,98)}=.915$, $p=.458$) in the difference-score analysis and together these results support the interpretation that brain-damaged patients were perceived by raters as generally depressed (See Figure 1).

Figure 1: Group x Emotion Identification Ratings

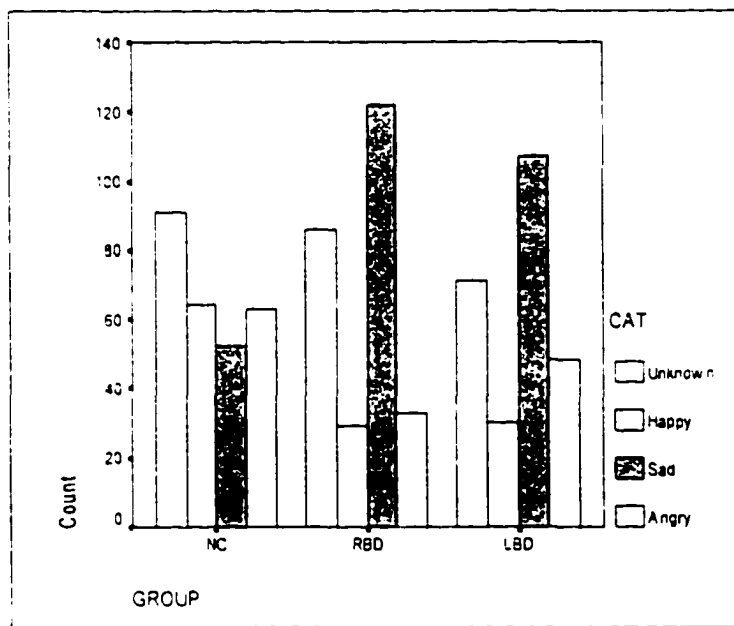


Furthermore, if a larger proportion of samples produced by brain-damaged patients were agreed by raters to sound sad compared to the normal controls, this result would be converging evidence for the interpretation that brain-damaged patients were perceived as depressed. Therefore, a Group x Category crosstabs analysis was conducted to assess this hypothesis and was significant ($\chi^2_{(6)}=59.170$, $p<.000$). Results conform to the interpretation that brain-damaged patients were perceived to sound depressed

resulting in a disproportionate number of samples to be categorized as sad relative to the rates of normal controls (See Figure 2).

Finally, a Mixed-effects analysis using lesion location, gender, and emotion as

Figure 2: Group x Category Frequency Counts



fixed effects and subject as a random effect was conducted on the identification ratings of all the samples to ascertain whether particular patients groups accounted for this effect. The lesion group x emotion effect was indeed significant

($F_{(12,2346)}=23.40, p<.0000$)

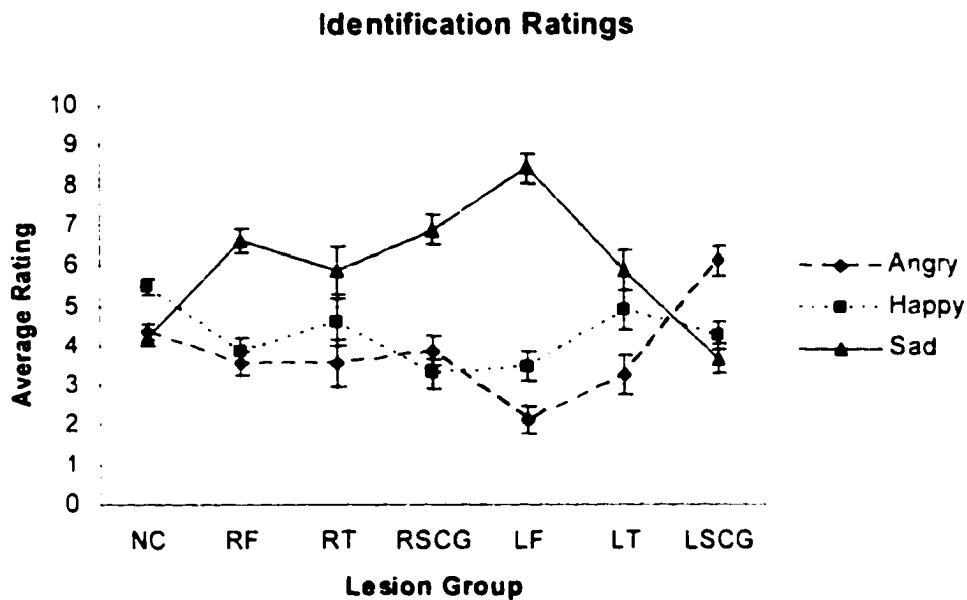
revealing that left-frontal patients

were the most impaired, followed by right-frontal and right-subcortical patients. Notably, left-subcortical patients were not significantly impaired, which is relevant for assessing cortical/subcortical hypotheses (See Figure 3).

Several possible explanations exist for the significant Emotion x Gender interaction effect found in the previous identification rating analysis. One explanation is that women are better at expressing happy and sad while men are better at expressing angry. Alternatively, women may have been perceived to sound more happy or sad in general, while men were perceived to sound angrier. In order to clarify this issue, identical analyses were performed using the difference-score between the emotional identification rating and neutral condition identification rating. Again, the Emotion x

Gender interaction effect was significant ($F_{(2,98)}=4.450, p=.014$). Post-hoc comparisons revealed that rater agreement for happy was no longer higher for women or higher for men on angry. However, for sad agreement rates, women's samples continued to be significantly higher than men. This supports the interpretation that women were generally perceived to sound happy, resulting in higher "happy" agreement rates in both

Figure 3: Lesion site x Emotion Identification Ratings



control and emotional conditions, while men were generally perceived to sound angrier, resulting in higher "angry" agreement rates in both control and emotional conditions. Additionally, these results also suggest that women were in fact better at conveying the emotion sad since this effect was maintained in the second analysis (See Figure 4).

Converging evidence for this interpretation would be found if more overall samples were deemed as happy for women and more samples deemed angry for men. A

Gender x Category crosstabs

analysis was conducted and was

also significant ($\chi^2_{(3)}=23.043$,

$p<.000$). Results conform to the

interpretation that women were

perceived to sound happy while

men were perceived to sound

angry. Consistent w/ results

stated above, this also suggested

that women express sad better

than men because there was not a

category frequency discrepancy

for sad by gender (See Figure 5).

Confidence ratings A

Group x Gender x Emotion

(Happy, Sad, Angry, Neutral)

repeated-measures ANOVA was examined on rater confidence scores of subjects' "best"

sample to assess differences in the level of rater confidence. Both the main effect of

Emotion ($F_{(3,147)}=2.785$, $p=.043$) and Gender ($F_{(1,49)}=15.792$, $p<.000$) were significant.

Additionally, the interaction effect of Gender x Emotion was also significant

($F_{(3,147)}=3.657$, $p=.014$), while the Group x Gender x Emotion interaction effect just

failed to reach significance ($F_{(6,147)}=2.135$, $p=.053$). All other effects were clearly not

Figure 4: Gender x Emotion Identification Ratings

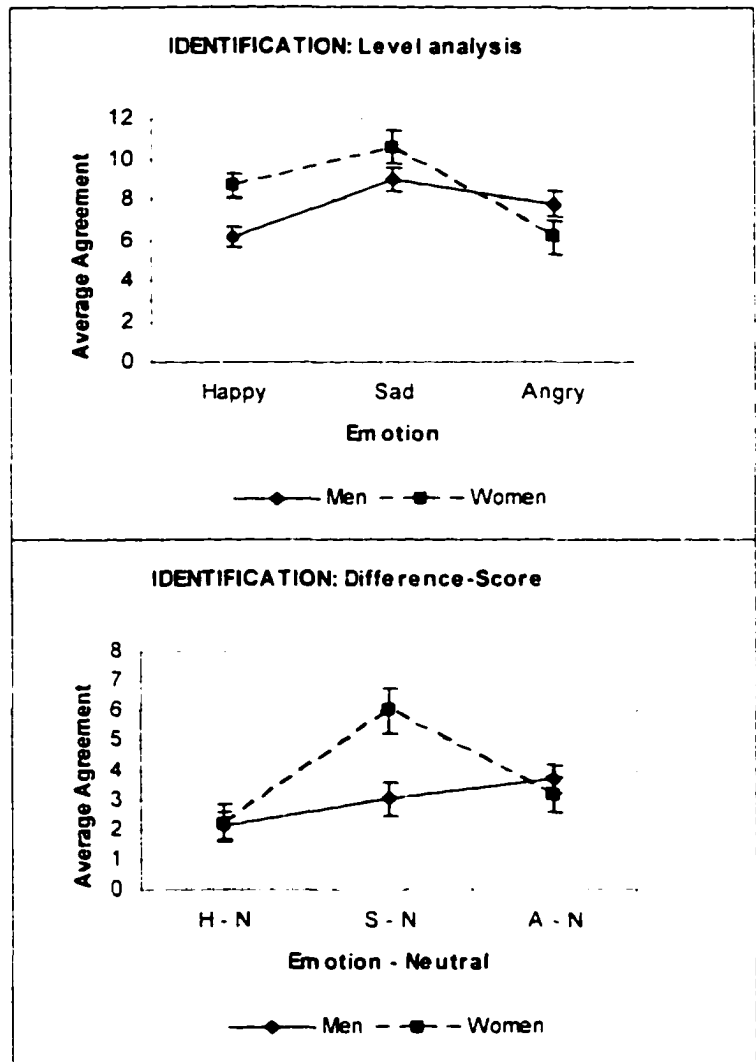
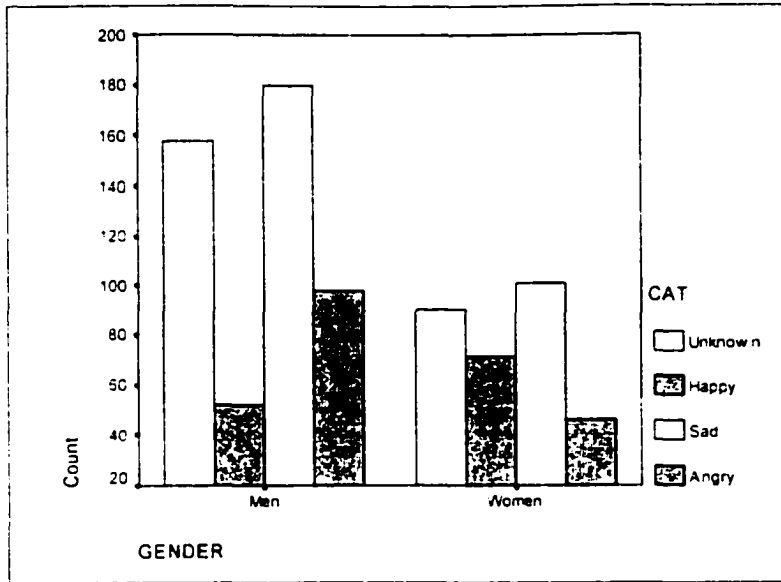


Figure 5: Gender x Category Frequency Counts



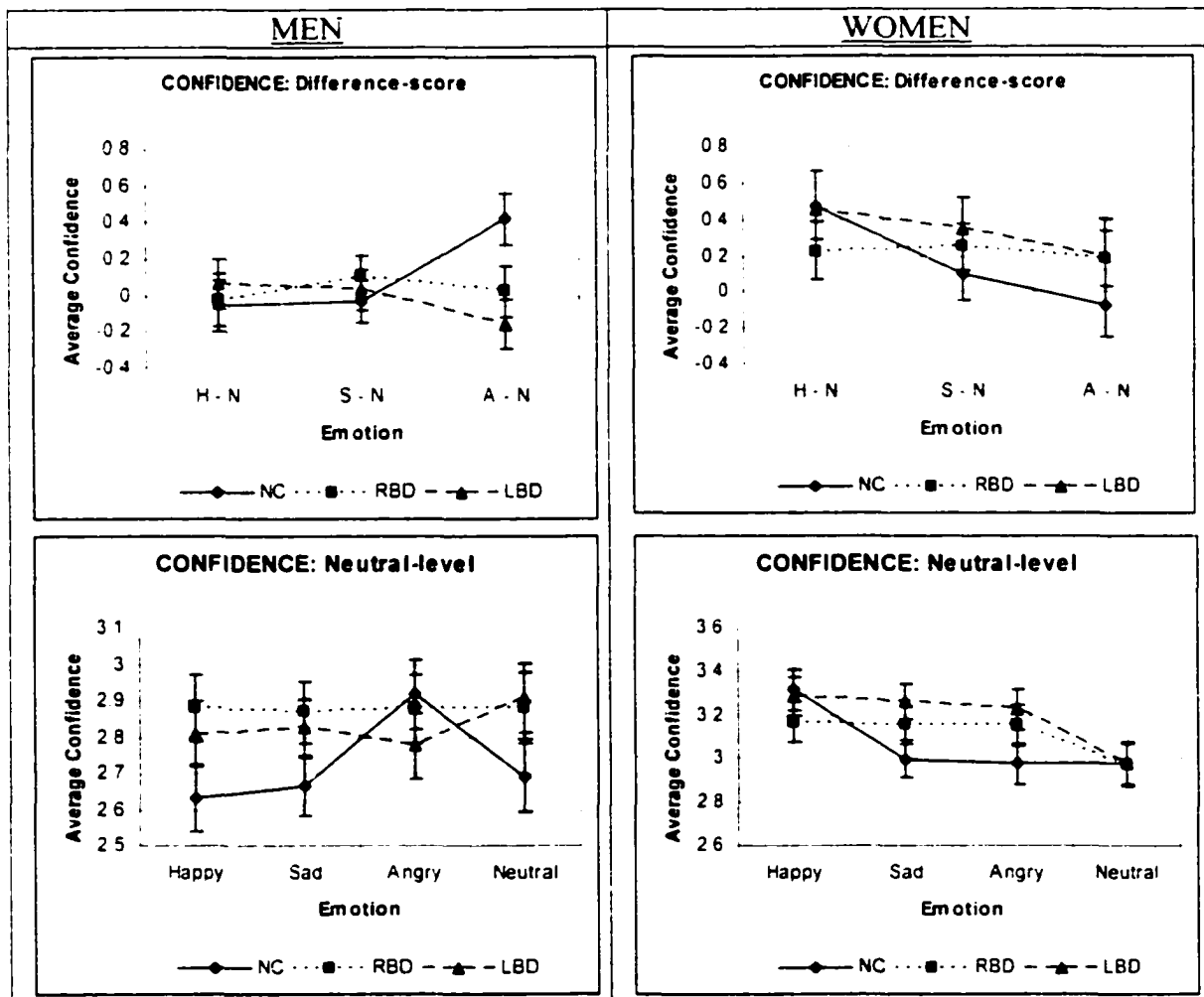
significant. Significant pairwise comparisons revealed that raters were more confident on samples deemed happy and sad compared to neutral samples, whereas there was not a significant increase in rater confidence for samples deemed angry over neutral

samples. Raters were also more confident rating happy and sad samples generated from women as opposed to men; this was not so for angry or neutral conditions.

A Group x Gender x Emotion (Happy – neutral, Sad – neutral, angry – neutral) repeated-measures ANOVA using differences scores of the “best” sample confidence ratings was also conducted to facilitate interpretation of the data. Again, the main effect of Gender ($F_{(1,49)}=4.582, p=.037$) and Gender x Emotion interaction effect ($F_{(2,98)}=3.371, p=.038$) were significant. Unlike the previous analysis, repeated contrasts revealed that raters were not more confident on sad samples from women over men. This discrepancy suggests that raters were generally more confident on samples generated by women but does not support the claim that women expressed sad better than men. However, raters continued to be significantly more confident on happy samples from women over men suggesting that women indeed prosodically expressed the emotion happy better than men.

Additionally, the Group x Gender x Emotion interaction effect ($F_{(4,98)}=2.550$, $p=.044$) reached significance in this analysis. Post hoc comparisons revealed that raters were most confident when rating angry samples from NC men and rating happy samples from women across groups. This pattern of data is in contrast with the previous analysis in which raters appeared to have greater confidence on samples produced by male BDs

Figure 6: Gender x Emotion Confidence Ratings



for happy, sad and neutral samples, but not angry. This suggests that NC men prosodically expressed the emotion angry better than men in other groups or women. On the other hand, raters had more confidence on happy samples from NC women than sad

and angry samples. The expression of BD women resulted in higher agreement levels than NC women in happy and sad conditions (See Figure 6).

Intensity ratings A Group x Gender x Emotion (Happy, Sad, Angry, Neutral) repeated-measures ANOVA was conducted to assess differences in perceived intensity level. A significant main effect of Emotion ($F_{(3,147)}=3.795, p=.012$) and Gender ($F_{(1,49)}=14.546, p<.000$) resulted. All other effects did not approach significance. Samples produced by women were perceived to be more emotionally intense than samples produced by males. Additionally, significant pairwise comparisons revealed that samples produced in happy, sad, and angry monologues were perceived as more emotionally intense than samples produced in neutral monologues. There were also significant differences among the emotions. A Group x Gender x Emotion (Happy – Neutral, Sad – Neutral, Angry – Neutral) repeated-measures ANOVA was also conducted on the “best” sample intensity ratings to assess differences in emotional expression. No significant effects were found, suggesting that women were judged to be more intense generally and did not have a significantly increased intensity during emotional monologues compared to men.

Subjects' Perspective - Acoustics

Introduction These analyses were conducted to find acoustic evidence to support the rating data as well as to potentially reveal effects not uncovered by human raters. Mean and standard deviation summaries of each parameter are available in Tables 13 and 14.

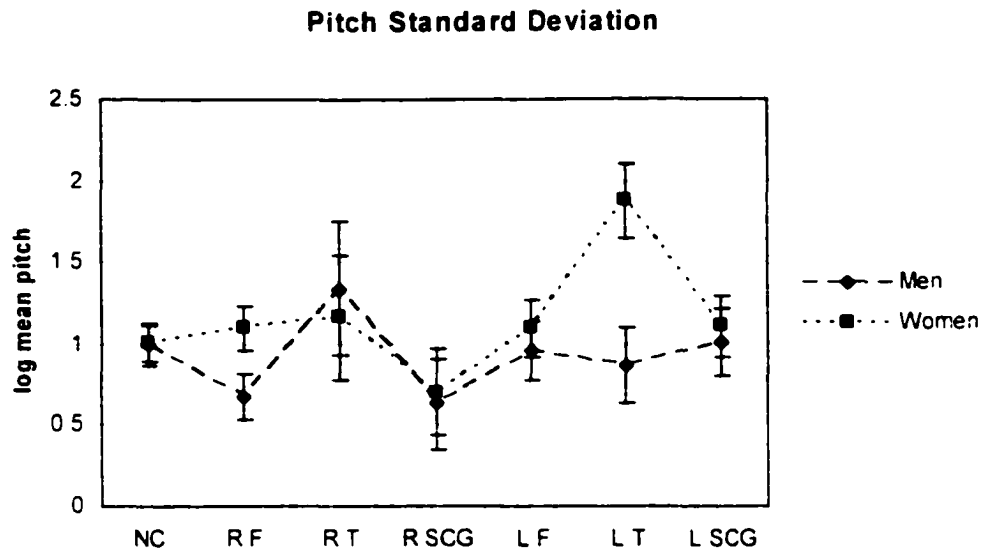
Pitch mean The neutral-level analysis revealed a significant effect of Gender ($F_{(2,49)}=53.836, p<.000$), such that women spoke in a significantly higher pitch than men. The effect of Gender was also significant in the difference-score analysis ($F_{(2,49)}=4.257,$

$p=.044$), which is meaningful because it suggests that women raised their pitch to express emotion to a greater extent than men did. A significant effect of Emotion (neutral-level: $F_{(3,147)}=8.522, p<.000$) revealed that happy and angry samples were higher in pitch than sad and neutral samples. This effect and data pattern maintains itself in the difference-score analysis ($F_{(2,98)}=10.4, p<.000$) suggesting that the increase in pitch is important to the prosodic expression of both happy and angry.

The Gender x Emotion effect failed to reach significance ($F_{(3,147)}=2.191, p<.092$) in the neutral-level analysis but reached significance in the difference-score ($F_{(2,98)}=3.888, p<.024$). Post hoc comparisons revealed that females significantly increased their mean pitch level relative to the control condition for happy and angry, but not for sad. Conversely, men did not show a significant pitch increase for any emotional condition relative to control levels. Together, this suggests that while the difference in pitch between men and women does not vary significantly across the emotional conditions, women increase their pitch over neutral levels to express happy and angry to a greater extent than men.

Pitch standard deviation The neutral-level analysis failed to produce any significant effects, suggesting the absence of categorical differences across groups or gender for this parameter. However, in the difference-score analysis, a significant Group x Gender interaction effect was found ($F_{(2,49)}=4.228, p=.02$). Repeated contrasts suggest that normal control men used a greater pitch range than normal control women, while the inverse was true for left-brain damaged patients. Greater pitch range is considered an important feature of expressing angry and therefore may support the interpretation of the rating data that normal control men sounded angry.

Figure 7: Pitch Standard Deviation by Lesion Site



In order to compare findings directly to the lesion site mixed effects analysis of the identification ratings, a similar mixed effects analysis was conducted on pitch standard deviation. The effect of lesion site was not statistically significant ($F_{(6,41)}=1.57$, $p=.178$). However, the data pattern mimics findings in the literature that right-frontal and right-subcortical patients appear to have restricted pitch range compared to other groups (See Figure 7). Nonetheless, these results do not directly support the cue-lateralization or right hemisphere hypotheses.

Pitch slope A significant effect of Emotion ($F_{(3,147)}=5.145$, $p=.002$) was found in the neutral-level analysis. Significant pairwise comparisons revealed that pitch acceleration was significantly faster in angry samples compared to both sad and neutral samples. Pitch acceleration in happy samples was not significantly different from any condition. This effect and data pattern maintains itself in the difference-score analysis

($F_{(2,98)}=3.322$, $p=.04$) supporting the interpretation that an increase in pitch acceleration is important in the emotional expression of angry.

Significant simple contrasts of the neutral-level Gender x Emotion ($F_{(3,147)}=3.691$, $p=.013$) interaction effect revealed that men used faster pitch acceleration than women in sad samples, and women had faster pitch acceleration in neutral samples. The relationship between men and women did not change significantly from the neutral condition for happy or angry samples. The Gender x Emotion interaction effect was also significant in the difference-score analysis ($F_{(2,98)}=4.190$, $p=.018$). Significant repeated contrasts revealed that women did not use pitch acceleration changes to facilitate the expression of happy or angry, but significantly slowed their pitch acceleration to express sad. Conversely, men significantly increased pitch acceleration to express angry and sad, but not happy. Taken together, these results suggest that to express sad, men increase pitch acceleration from neutral levels while women decrease it. Additionally, although pitch acceleration levels to express angry are not significantly different in level between men and women, men significantly increased pitch acceleration from neutral levels to express angry and women do not.

Pitch regression 1st half sample The Gender x Emotion interaction effect in the neutral-level analysis was significant ($F_{(3,147)}=3.339$, $p=.021$) revealing women maintain a steady pitch level in happy samples. This was significantly higher than the declining pitch regression observed in sad and angry samples from women; men's pitch regression in angry samples was significantly less negative than women's. The Gender x Emotion effect ($F_{(2,98)}=4.336$, $p=.016$) in the difference-score analysis told a similar story, as contrasts revealed that women relative to men, had a significantly greater pitch regression change in happy but that men increased their pitch regression in angry and women's

declined. Together these results suggest that women use a steady pitch level to express happy and men use a steady pitch to express angry. Pitch regression 2nd half sample Both the main effect of Emotion ($F_{(3,147)}=3.339$, $p=.021$) and Gender ($F_{(3,147)}=3.339$, $p=.021$) were significant in the neutral-level analysis, while only the main effect of Emotion ($F_{(2,98)}=4.7$, $p=.011$) was significant in the difference score analysis. The effect of Gender was clearly not significant ($F_{(1,49)}=.288$, $p=.594$). These results were modified by the significant interaction effect of Gender x Emotion in both the neutral-level ($F_{(3,147)}=4.514$, $p=.005$) and difference-score ($F_{(2,98)}=4.456$, $p=.014$) analyses. Simple contrasts of the neutral-level analysis revealed that the difference in declining pitch between women and men was greater for happy samples and smaller for sad samples compared to neutral levels. Repeated contrasts of the difference-score analysis show that women dropped their pitch significantly more than men in happy samples but pitch changes from neutral levels were similar for the other emotions. Together, these results suggest women's pitch declines in the 2nd half of samples more than men, but this drop reflects a difference in the expression of emotion in happy samples only.

Pitch regression entire sample The main effect of Group ($F_{(2,49)}=3.519$, $p=.037$) was significant in the neutral-level analysis but failed to reach significance in the difference-score analysis ($F_{(2,49)}=1.151$, $p=.325$). Significant pairwise comparisons from the first analysis revealed that the pitch of LBDs declined throughout the sample at a significantly greater rate than NCs. However, given the absence of a significant difference-score effect, this difference appears to be one of level rather than a function of expressing emotion.

There were also significant main effects of Emotion ($F_{(3,147)}=4.325$, $p=.006$) and Gender ($F_{(1,49)}=26.49$, $p<.000$) in the neutral level analysis, but only a significant effect of

Emotion ($F_{(2,98)}=5.643$, $p=.005$) in the difference-score analysis. These effects were modified by significant Gender x Emotion interaction effects (neutral-level: $F_{(3,147)}=3.732$, $p=.013$; difference-score: $F_{(2,98)}=4.138$, $p=.019$). Significant repeated contrasts of the neutral-level effect revealed that while men were steady, the difference between men and women was greater for happy and anger samples compared to sad and neutral samples. Contrasts of the difference-score effect revealed that women's pitch declined significantly more from neutral levels in happy and angry samples compared to men's, but rose significantly more than men's in sad samples. Together, these results suggest that women use gradual changes in pitch level to convey emotion to a greater extent than men, such that women's pitch falls from neutral levels for happy and angry and rising from neutral levels for sad.

dB standard deviation There was only a significant effect of Emotion in the neutral-level analysis ($F_{(3,147)}=2.941$, $p=.035$). Pairwise comparisons revealed that happy and angry had significantly more volume variation than sad, but none were significantly different from neutral levels. This effect was also significant in the difference-score analysis ($F_{(2,98)}=2.941$, $p=.035$). Pairwise comparisons supported the same conclusion, i.e., that volume variation is increased to expressed happy and angry and decreased to express sad. The absence of Gender or Group effects suggests that all sub-groups were able to modulate volume to similar effect.

Speech rate There was a significant main effect of Group ($F_{(2,49)}=4.469$, $p=.016$) in the neutral-level analysis. Pairwise comparisons revealed that LBDs spoke significantly slower than NCs. The Group effect was also significant in the difference-score analysis ($F_{(2,49)}=4.453$, $p=.017$). Pairwise comparisons suggested that the speech rate of NCs increased in emotional conditions, and fell for LBDs. RBDs showed

negligible speech rate changes in emotional conditions and was not significantly different from either group.

The Group x Emotion interaction effect was significant ($F_{(6,147)}=2.223, p=.043$) in the neutral level analysis, but failed to reach significance in the difference-score analysis ($F_{(4,98)}=1.107, p=.358$). Repeated contrasts revealed that speech rate of LBDs significantly declined in all emotional conditions relative to neutral levels, while NCs speech rates increased. The absence of a significant Group x Emotion effect in the difference score analysis suggests that this effect is a function of the difference in conditions between neutral and emotional conditions and not one of the expression of emotion.

Percent of sample with no voicing

Consistent with the significant Group effect of speech rate, the main effect of Group ($F_{(2,49)}=2.805, p=.07$) approached significance in the difference-score analyses; the trend was for LBDs to demonstrate less pausing in the emotional conditions relative to neutral conditions compared to normal controls. This is meaningful as it supports the Group speech rate finding and the interpretation that the tempo with which LBDs spoke in emotional conditions was categorically different than in the neutral conditions.

The Gender x Emotion interaction effect was significant ($F_{(3,147)}=3.499, p=.017$) in the neutral-level analysis. Repeated contrasts revealed that women had more silence than men in sad samples. This effect was also significant in the difference-score analysis ($F_{(2,98)}=3.773, p=.026$). Repeated contrasts revealed that women had significantly more silence time in sad samples relative to neutral levels, and men had lower levels of silence in sad and angry samples relative to neutral levels. Together, these results suggest women use silence to express sad whereas men reduce silence. Further, although men

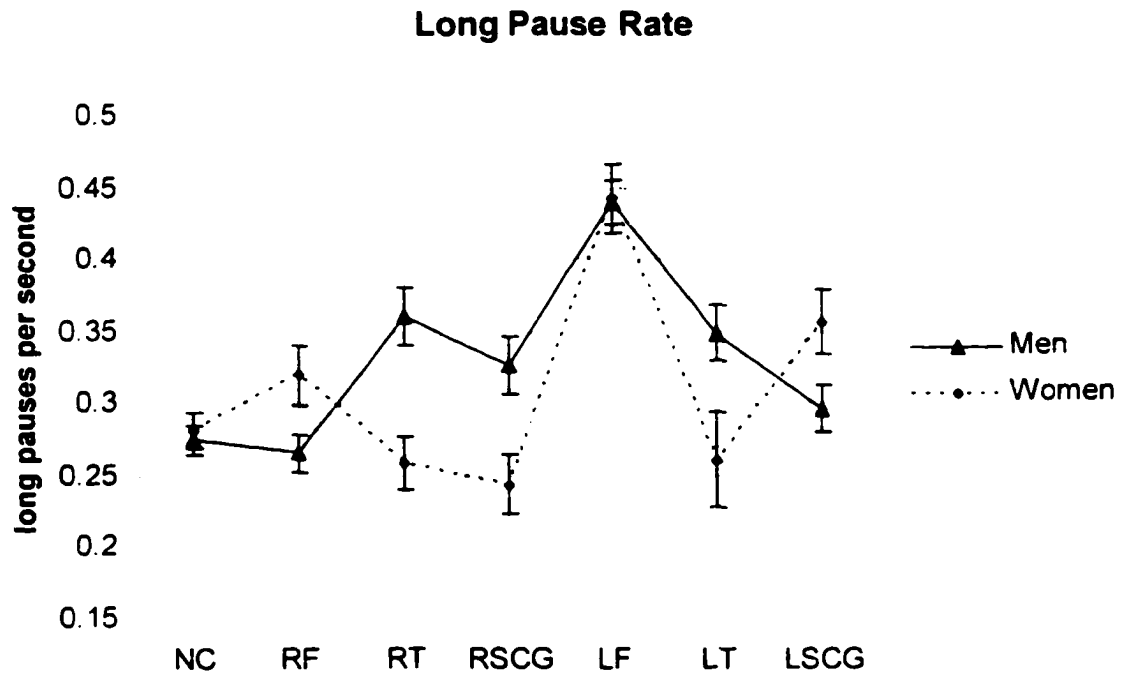
and women did not differ in level of silence in angry samples, men showed a significant decline in silence levels from neutral levels to express anger and women did not significantly change silence levels from neutral levels.

Short Pause Rate No significant effects were found in either analysis.

Long Pause Rate There was a significant Group effect in the neutral-level analysis ($F_{(2,49)}=3.805$, $p=.029$). Pairwise comparisons revealed that LBDs had a significantly higher long pause rate than either RBDs or NCs. The Group effect was not significant in the difference score analysis ($F_{(2,49)}=.848$, $p=.435$) suggesting that the increase rate of long pauses was not a function of expressing emotion.

The Gender x Emotion interaction effect was significant in the neutral-level analysis ($F_{(3,147)}=3.672$, $p=.014$). Simple contrasts revealed that women had a higher large pause rate in the sad condition compare to neutral levels, and men had a lower pause rate for sad samples relative to neutral levels. This effect was also significant in the difference-score analysis ($F_{(2,98)}=3.228$, $p=.044$), repeated contrasts revealed that women had increased the long pause rate in sad samples relative to neutral levels and men lowered their long pause rate in sad samples relative to neutral levels. Men also lowered their long pause rate in angry samples relative to neutral levels compared to women who showed a slight increase in pause rate. Together, these results suggest women used increased long pauses to express sad and men decreased their long pauses rate to express sad and angry.

Figure 8: Long Pause Rate by Group



Raters' Perspective – Acoustic to Category relationships

Introduction The goal of the following analyses was to determine if the acoustic characteristics of the speech samples can account for or predict rater judgments of emotional identification, confidence, and intensity. Unlike the previous analyses which were based on the subjects' emotional intent, these analyses are grounded on the raters' perception of the samples. For example, a subject may have intended to convey the emotion angry in a sample, but raters perceived the speech as sad and therefore the sample was treated as an example of sad expressing for these analyses. The process of giving a "category" to each sample is described in detail under data preparation in the Methods section.

Category

Introduction Separate mixed-effects models using each acoustic parameter as a response variable were conducted. Results are broken down by the response variable rather than the fixed effects (Category x Strength x Gender).

Pitch mean Analysis revealed a significant effect of category ($F_{(3,554)}=22.15$, $p<.0001$), gender ($F_{(1,53)}=52.34$, $p<.0001$), and the category by strength interaction ($F_{(6,554)}=3.21$, $p=.0042$). The Akaike Information Criterion (AIC) goodness-of-fit measure was (-1408.352) and the Bayesian Information Criterion (BIC) was (-1273.532). Significant post-hoc comparisons revealed that mean pitch was highest for happy, followed by angry and that the mean pitch for sad is significantly lower than for the unknown category. The interaction of category and strength shows that although the mean pitch was high for all levels of happy, the mean pitch of angry strength level 3 is equal to happy strength 3; it was also significantly lower at strength levels 1 and 2, indicating that fewer raters thought the sample sounded angry. Finally, women had a significantly higher mean pitch than men, which is consistent with previous studies.

Pitch standard deviation A significant main effect of category ($F_{(3,554)}=5.86$, $p=.0006$) was revealed, as were significant interaction effects of category by gender ($F_{(3,554)}=4.3368$, $p=.0049$) and category by strength ($F_{(6,554)}=2.3795$, $p=.0280$). The AIC goodness-of-fit measure was (540.54) and the BIC was (675.36). Significant post-hoc comparisons revealed that the pitch standard deviation of happy was significantly greater than the other three categories. Furthermore, as agreement increased from strength level 2 to 3, pitch standard deviation increased for samples identified angry but declined for sad. Finally, samples produced by men deemed angry, and to a lesser extent happy, had a

significantly smaller pitch standard deviation compared to samples deemed angry and happy produced by women.

Pitch slope There were no significant effects for pitch slope.

Pitch regression 1st half sample Results revealed a significant main effect of gender ($F_{(1,53)}=4.89$, $p=.0313$) and significant interaction effects of category by gender ($F_{(3,554)}=3.42$, $p=.0169$) and gender by strength ($F_{(2,554)}=4.33$, $p=.0136$). The AIC goodness-of-fit measure was (-2533.68) and the BIC was (-2398.86). Significant post hoc comparisons revealed that both men and women had a falling pitch in the first half of samples, but women's pitch fell significantly more than men's. Male pitch slope fell less in samples rated as sad and angry than in other samples. Finally, pitch regression fell less in samples produced by men, but not women, as more raters agreed on the emotion in the sample.

Pitch regression 2nd half sample A significant main effect of gender ($F_{(1,53)}=17.84$, $p<.0001$) was found, while the main effect of category ($F_{(3,554)}=2.49$, $p=.0592$) and category by gender ($F_{(3,554)}=2.35$, $p=.0715$) just failed to reach significance. The AIC goodness-of-fit measure was (-4460.42) and the BIC was (-4325.61). Examination of the main effect of gender shows that women's pitch fell significantly more than men's in the second half of the sample.

Pitch regression entire sample Results revealed a significant main effect of category ($F_{(3,554)}=4.54$, $p=.0037$) and gender ($F_{(1,53)}=23.14$, $p<.0001$). The category by gender interaction was the only other effect to approach significance ($F_{(3,554)}=2.35$, $p=.0712$). The AIC goodness-of-fit measure was (-5334.94) and the BIC was (-5200.11). Post hoc comparisons revealed that the downward slope of pitch over the course of samples was significantly less pronounced for sad samples than happy or angry samples.

Furthermore, the pitch of females fell significantly more than men's over the entirety of samples.

Table 13: Group x Emotion effects - Best Sample Pitch Acoustics

		Happy Mean \pm (SD)	Sad Mean \pm (SD)	Angry Mean \pm (SD)	Neutral Mean \pm (SD)
F ₀ mean	NC	132.76 \pm (49.59)	124.66 \pm (32.21)	135.57 \pm (43.67)	128.25 \pm (39.68)
	RBD	136.35 \pm (36.14)	131.95 \pm (36.37)	137.45 \pm (35.64)	129.84 \pm (33.35)
	LBD	136.35 \pm (39.95)	121.72 \pm (36.64)	138.88 \pm (44.39)	131.97 \pm (34.79)
F ₀ Std Dev	NC	2.90 \pm (1.04)	2.92 \pm (1.12)	3.39 \pm (1.21)	2.78 \pm (1.01)
	RBD	2.64 \pm (1.34)	2.80 \pm (1.64)	2.74 \pm (1.33)	2.49 \pm (1.27)
	LBD	3.54 \pm (1.71)	2.81 \pm (1.69)	3.14 \pm (1.54)	3.38 \pm (1.39)
F ₀ Slope	NC	20.51 \pm (5.24)	22.17 \pm (7.37)	24.25 \pm (6.57)	22.67 \pm (10.47)
	RBD	23.58 \pm (7.19)	22.32 \pm (7.94)	24.62 \pm (8.54)	22.26 \pm (6.27)
	LBD	20.13 \pm (7.25)	20.80 \pm (10.96)	21.69 \pm (8.74)	17.73 \pm (7.86)
F ₀ reg. 1st	NC	-0.007 \pm (0.05)	-0.036 \pm (0.04)	-0.007 \pm (0.05)	-0.019 \pm (0.03)
	RBD	-0.005 \pm (0.04)	-0.015 \pm (0.03)	-0.013 \pm (0.05)	-0.012 \pm (0.02)
	LBD	-0.013 \pm (0.07)	-0.008 \pm (0.03)	-0.015 \pm (0.04)	-0.017 \pm (0.03)
F ₀ reg. 2nd	NC	0.001 \pm (0.05)	0.011 \pm (0.03)	0.003 \pm (0.05)	0.008 \pm (0.04)
	RBD	-0.012 \pm (0.03)	-0.011 \pm (0.02)	-0.018 \pm (0.03)	-0.009 \pm (0.02)
	LBD	-0.016 \pm (0.03)	-0.007 \pm (0.03)	-0.012 \pm (0.03)	-0.012 \pm (0.03)
F ₀ reg. entire	NC	0.002 \pm (0.02)	0.005 \pm (0.02)	0.001 \pm (0.03)	0.002 \pm (0.02)
	RBD	-0.010 \pm (0.02)	-0.007 \pm (0.01)	-0.010 \pm (0.02)	-0.006 \pm (0.01)
	LBD	-0.011 \pm (0.02)	-0.005 \pm (0.02)	-0.009 \pm (0.02)	-0.008 \pm (0.02)

dB standard deviation A significant main effect of category ($F_{(3,554)}=12.64$,

$p<.0001$) was found in addition to a significant category by gender interaction effect ($F_{(3,554)}=3.112$, $p=.0260$). The AIC goodness-of-fit measure was (-870.68) and the BIC was (-735.86). Examination of post hoc comparisons revealed that samples judged sad by raters possessed the smallest dB standard deviation while happy samples had a significantly greater dB standard deviation than the other categories. Samples judged angry had a significantly smaller dB standard deviation than happy, but greater than unknown and sad samples (sad < unknown < angry < happy). Furthermore, this data suggests that females increase the range of loudness when expressing anger and men decrease the range.

		Happy	Sad	Angry	Neutral
		Mean \pm (SD)	Mean \pm (SD)	Mean \pm (SD)	Mean \pm (SD)
dB mean	NC	69.51 \pm (1.72)	70.15 \pm (2.01)	69.54 \pm (1.65)	69.15 \pm (2.41)
	RBD	68.85 \pm (1.68)	68.74 \pm (2.11)	68.85 \pm (2.20)	69.07 \pm (1.65)
	LBD	68.30 \pm (2.02)	69.07 \pm (1.63)	68.21 \pm (1.76)	68.54 \pm (1.81)
dB Std Dev	NC	8.87 \pm (1.89)	8.52 \pm (1.90)	8.91 \pm (1.53)	8.89 \pm (2.73)
	RBD	9.10 \pm (2.00)	8.81 \pm (2.23)	9.25 \pm (2.35)	8.88 \pm (2.03)
	LBD	9.41 \pm (1.97)	8.15 \pm (1.71)	9.39 \pm (2.06)	9.07 \pm (2.18)
Speech rate	NC	417.79 \pm (83.21)	406.56 \pm (59.99)	462.33 \pm (102.76)	389.60 \pm (72.60)
	RBD	366.31 \pm (80.00)	406.37 \pm (114.28)	390.39 \pm (85.96)	390.89 \pm (76.06)
	LBD	343.91 \pm (100.80)	335.07 \pm (88.93)	357.04 \pm (64.98)	390.64 \pm (83.43)
% no voice	NC	46% \pm (6%)	43% \pm (9%)	44% \pm (9%)	41% \pm (11%)
	RBD	42% \pm (9%)	45% \pm (13%)	41% \pm (9%)	44% \pm (8%)
	LBD	44% \pm (10%)	46% \pm (8%)	45% \pm (6%)	51% \pm (9%)
Short Pause Rate	NC	0.252 \pm (0.18)	0.270 \pm (0.16)	0.268 \pm (0.14)	0.218 \pm (0.14)
	RBD	0.248 \pm (0.14)	0.235 \pm (0.12)	0.231 \pm (0.16)	0.251 \pm (0.11)
	LBD	0.248 \pm (0.14)	0.224 \pm (0.12)	0.241 \pm (0.11)	0.314 \pm (0.17)
Long Pause Rate	NC	0.327 \pm (0.12)	0.266 \pm (0.16)	0.272 \pm (0.13)	0.266 \pm (0.13)
	RBD	0.287 \pm (0.12)	0.328 \pm (0.18)	0.272 \pm (0.12)	0.295 \pm (0.07)
	LBD	0.348 \pm (0.14)	0.388 \pm (0.10)	0.356 \pm (0.11)	0.393 \pm (0.10)

Speech rate Only the main effect of category ($F_{(3,554)}=4.035$, $p=.0074$) was significant. The AIC goodness-of-fit measure was (9121.505) and the BIC was (9256.33). Planned comparisons revealed that the speech rate of samples judged happy was significantly faster than for the other categories, and that sad samples had the slowest speech rate. Samples judged angry had a speech rate significantly faster than sad, but significantly slower than happy; samples in the unknown category had an average speech rate between angry and sad (Happy > Angry > unknown > sad).

Percent of sample with no voicing Only the interaction effect of category by gender was significant ($F_{(3,554)}=4.02$, $p=.0076$). The AIC goodness-of-fit measure was (-1653.73) and the BIC was (-1518.91). Significant post hoc comparisons revealed that females had less no voicing time in samples judged happy and angry than did men, but had more time with no voicing in samples judged sad.

Table 15: Gender x Emotion effects - Best Sample Acoustics

		Happy		Sad		Angry		Neutral	
		Mean	± (SD)	Mean	± (SD)	Mean	± (SD)	Mean	± (SD)
F ₀ mean	M	110.62	± (17.98)	107.12	± (19.91)	114.78	± (20.11)	110.26	± (19.29)
	F	173.97	± (40.62)	156.40	± (33.29)	172.91	± (41.30)	161.24	± (34.32)
F ₀ Std Dev	M	2.85	± (1.22)	2.60	± (1.14)	2.89	± (1.12)	2.64	± (1.22)
	F	3.28	± (1.68)	3.24	± (1.88)	3.40	± (1.68)	3.24	± (1.31)
F ₀ Acceleration	M	20.81	± (6.78)	23.33	± (10.71)	24.23	± (9.66)	20.55	± (10.14)
	F	22.40	± (6.69)	19.34	± (3.48)	22.47	± (4.40)	21.56	± (5.56)
F ₀ slope 1st	M	-0.017	± (0.05)	-0.012	± (0.04)	0.002	± (0.04)	-0.01	± (0.02)
	F	0.005	± (0.06)	-0.033	± (0.03)	-0.033	± (0.04)	-0.02	± (0.04)
F ₀ slope 2nd	M	0.007	± (0.02)	0.009	± (0.02)	0.004	± (0.02)	0.01	± (0.02)
	F	-0.035	± (0.04)	-0.020	± (0.03)	-0.029	± (0.05)	-0.02	± (0.04)
F ₀ slope entire	M	0.003	± (0.01)	0.004	± (0.01)	0.004	± (0.01)	0.00	± (0.01)
	F	-0.022	± (0.02)	-0.013	± (0.02)	-0.022	± (0.03)	-0.02	± (0.02)
dB mean	M	69.01	± (1.99)	69.72	± (2.15)	69.04	± (2.09)	68.83	± (1.98)
	F	68.72	± (1.66)	68.69	± (1.60)	68.63	± (1.71)	69.08	± (2.04)
dB Std Dev	M	8.95	± (1.97)	8.35	± (2.19)	9.10	± (2.06)	9.09	± (2.49)
	F	9.41	± (1.93)	8.73	± (1.57)	9.31	± (1.94)	8.72	± (2.05)
Speech rate	M	395.25	± (89.90)	381.67	± (83.46)	409.20	± (103.52)	399.30	± (78.61)
	F	346.96	± (91.30)	386.41	± (113.96)	395.94	± (84.38)	376.24	± (73.09)
% no voice	M	44%	± (9%)	42%	± (10%)	42%	± (8%)	46%	± (11%)
	F	43%	± (8%)	48%	± (8%)	44%	± (8%)	43%	± (7%)
Short Pause Rate	M	0.262	± (0.172)	0.226	± (0.153)	0.229	± (0.139)	0.265	± (0.163)
	F	0.228	± (0.116)	0.270	± (0.108)	0.273	± (0.136)	0.252	± (0.102)
Long Pause Rate	M	0.325	± (0.131)	0.290	± (0.160)	0.301	± (0.120)	0.336	± (0.117)
	F	0.312	± (0.122)	0.382	± (0.142)	0.296	± (0.141)	0.286	± (0.112)

Short Pauses The category by gender interaction effect was significant

($F_{(3,554)}=2.68$, $p=.0462$), as was the gender by strength interaction ($F_{(2,554)}=3.17$,

$p=.0428$). The AIC goodness-of-fit measure was (3440.95) and the BIC was (3575.77).

Significant post hoc comparisons revealed that men had significantly more short pauses than women when expressing happy, but produced significantly less short pauses than females for samples judged as angry. There was not a significant difference between the numbers of pauses for "sad" or "unknown" samples. Women had significantly more short pauses for strength levels one and two than men. Men had a sharp increase in short pauses associated with strength level 3 while the number of short pauses declined for women at this strength.

Long Pauses No significant effects were found for the number of long pauses.

Intensity

Introduction Confidence and intensity ratings were used as response variables in separate linear mixed-effects models to test whether differences in acoustic parameter are significant predictors of rater confidence and intensity judgments.

Angry model Mean pitch was the biggest predictor of anger intensity ratings ($F_{(1,450)}=54.81, p<.0001$) followed by pitch standard deviation ($F_{(1,450)}=15.88, p=.0001$) and volume standard deviation ($F_{(1,450)}=12.37, p=.0005$). Though not significant predictors themselves, pitch regression measures improved the fit of the model as reflected in significant decreases in goodness-of-fit measures (See table 15). Neither laterality of brain damaged group nor gender added significantly to the model. Goodness-of-fit measures were AIC (1618.69), BIC (1677.07) and logLik (-796.3463). Examination of the slope coefficients revealed that the estimated slope of mean pitch (.762) was significant ($t_{(450)}=3.159, p=.0017$) suggesting that as mean pitch increased, the rated intensity of angry also increased. The estimated slope of pitch standard deviation (.379) was also significant ($t_{(450)}=4.278, p<.0001$) revealing a positive relationship between pitch standard deviation and rated intensity of angry. Finally, the estimated slope of volume standard deviation (.574) was also significant ($t_{(450)}=3.477, p=.0006$) and positively related to rated intensity of angry.

Happy model Mean pitch was the biggest predictor of happy intensity ratings ($F_{(1,523)}=46.71, p<.0001$) followed by pitch standard deviation ($F_{(1,523)}=21.60, p<.0001$) and volume standard deviation ($F_{(1,523)}=16.87, p<.0001$). Additionally, pitch slope ($F_{(1,523)}=3.181, p=.0751$) and speech rate ($F_{(1,523)}=3.82, p=.0502$) both approached significance as predictors and did significantly increase the fit of the model. As with

angry. pitch regression measures improved the fit of the model but were not significant predictors of happy intensity ratings themselves (See Table 15). Neither brain damaged group or gender added significantly to the model. Goodness-of-fit measures were AIC (1805.789), BIC (1874.90), and logLik (-887.8943). Examination of the slope coefficients revealed that the estimated slope of mean pitch (.957) was significant ($t_{(523)}=3.485$, $p=.0005$) suggesting that as mean pitch increased the rated intensity of happy also increased. The estimated slope of pitch standard deviation (.329) was also positively related to rated intensity of anger ($t_{(523)}=3.884$, $p=.0001$). The estimated slope of volume standard deviation (.736) was also positively related to rated intensity of happy ($t_{(523)}=4.306$, $p<.0001$). It is also worth noting that the estimated slope of pitch slope was positive going (.358) and reached significance ($t_{(523)}=2.745$, $p=.0063$) and the estimated slope of speech rate (.0006) failed to reach significance ($t_{(523)}=1.857$, $p=.0639$).

Sad model The number of long pauses (pauses greater than 0.4 seconds) was the biggest predictor of sad intensity ratings ($F_{(1,487)}=12.002$, $p=.0006$) followed by mean pitch ($F_{(1,487)}=6.765$, $p=.0096$) and pitch regression during the 1st half of the sample ($F_{(1,487)}=5.663$, $p=.0177$). Additionally, the pitch regression of the entire sample and volume standard deviation significantly improved the fit of the model (See Table 15). Again, neither brain damaged group or gender contributed significantly to the model. Goodness-of-fit measures were AIC (1423.82), BIC (1478.52), and logLik (-699.9112). Examination of the slope coefficients revealed that long pauses (.031) were positively related to rated intensity of sad ($t_{(487)}=3.046$, $p=.0024$) as was pitch regression during the first half of the sample (1.42; $t_{(487)}=2.456$, $p=.0144$). Although considered a significant predictor based on ANOVA results, the estimated slope of mean pitch rate (.295) failed to reach significance ($t_{(487)}=1.24$, $p=.2140$).

Table 16: Mixed-effects model results - Acoustical parameter predictors of intensity

	HAPPY INTENSITY			SAD INTENSITY			ANGRY INTENSITY		
	<u>AIC</u>	<u>BIC</u>	<u>Loglik</u>	<u>AIC</u>	<u>BIC</u>	<u>Loglik</u>	<u>AIC</u>	<u>BIC</u>	<u>Loglik</u>
	1805.8	1874.9	-887.9	1423.8	1478.5	-699.9	1618.7	1677.1	-796.3
	ANOVA <i>F</i> (<i>p</i> -value)	SLOPE Est(<i>p</i> -value)	ANOVA <i>F</i> (<i>p</i> -value)	SLOPE Est(<i>p</i> -value)	ANOVA <i>F</i> (<i>p</i> -value)	SLOPE Est(<i>p</i> -value)			
<u>Pitch</u>									
Mean	46.71(.0001)	0.95(.0005)	6.77(.0096)	0.29(.2140)	54.81(.0001)	0.76(.0017)			
Std Dev	21.60(.0001)	0.32(.0001)			15.88(.0001)	0.37(.0001)			
Slope									
Regr. 1 st half			5.66(.0177)	1.42(.0144)					
Regr. 2 nd half									
Regr. entire									
<u>Volume</u>									
dB mean									
dB Std Dev	16.85(.0001)	0.73(.0001)			12.37(.0005)	0.57(.0006)			
<u>Tempo</u>									
Speech rate									
% no voice									
Short Pauses									
Long Pauses			12.00(.0006)	0.03(.0024)					

Discussion

Spontaneous speech samples were recorded from RBD, LBD, and NC subjects, low-pass filtered, and presented to naïve raters. Each rater decided whether the speaker in each sample sounded happy, sad, or angry (i.e., Identification rating), how confident they were in their choice (i.e., Confidence rating), and how intensely they believed the speaker was expressing the emotion they chose (i.e., Intensity rating). Higher rater agreement was interpreted to mean that the speech sample was a clearer example of the emotion chosen. In addition, each sample underwent acoustical analysis generating several measurements in three basic categories. Pitch cues included the average pitch level of the speaker's voice (i.e., pitch mean), the range of pitch used (i.e., pitch standard

deviation), how quickly the speaker's voice changed pitch (i.e., pitch slope), and the overall rising and falling of pitch over the sample length (i.e., pitch regression). Duration cues evaluate periods of articulation and silence (Pittam & Scherer, 1993) and included how quickly the speaker spoke (i.e., speech rate), proportion of silence contained in the sample (i.e., percent no voice), and the length of silences used (i.e., number of short pauses [0.2 – 0.4 seconds]; number of long pauses [> 0.4]; the number of short and long pauses corrected for sample time [i.e., short and long pause rate]). Lastly, the amplitude range of the speaker's voice was also measured (i.e., volume standard deviation).

The array of analyses generated from this data supported conclusions in three areas: 1) the acoustical parameters used to communicate happy, sad, and angry; 2) differential use of acoustics by men and women to express emotion; and 3) the brain systems important in the expression of prosodic emotion. Regarding the acoustic profiles of each emotion, results revealed that pitch mean, pitch standard deviation, volume standard deviation, and speech rate were all highest for the expression of happy and lowest for sad. The expression of angry was also associated with high pitch mean and standard deviation levels, however, only at the highest agreement level. Further, volume standard deviation was less variable and speech rate was not as fast in samples identified as angry relative to happy samples. In general, raters based their decision regarding emotional intensity on the level of these acoustics.

Analyses of gender by brain-damaged group were largely not significant and therefore conclusions regarding the significant gender effects are made across group. These results strongly suggest that men and women use pitch and duration cues differently to express emotion. Women, and not men, significantly increased their mean pitch from neutral levels to express happy and angry. Similarly, women used gradual

changes in pitch regression over the course of speaking (e.g., larger fall of pitch for happy and angry compared to neutral and less fall of pitch for sad compared to neutral). Men used a greater increase in pitch standard deviation to express emotion than women: men also used increases in pitch slope from neutral levels to express angry and women did not. Men, and not women, decreased the amount of silence time and long pauses from neutral levels to express angry.

Further, multiple analyses suggested that women tend to sound happy and men tend to sound angry when speaking neutrally. Samples produced by women were judged to sound happy at a higher frequency and rater agreement level while samples by men were judged angry at a higher frequency and rater agreement level. The identification rating data showed elevated agreement for samples judged happy for women and angry for men. A significantly greater proportion of samples produced by men were agreed to sound angry and those produced by women to sound happy. Confidence rating results also revealed that raters were most confident when rating angry samples from normal control men and rating happy samples from women from all groups. NC men also appear to express the emotion angry better prosodically than other groups including women and raters also declared more confidence on happy samples produced by NC women than sad and angry samples.

Finally, the identification rating data by lesion site revealed that left-frontal patients were the most impaired group followed by right-frontal and right-subcortical patients. Importantly, left subcortical patients did not appear to be impaired based on the identification ratings. Left-frontal patients had difficulty with duration cues, specifically having significantly more silence time and a greater number of long pauses in their samples, and in general, LBD patients spoke at a significantly slower rate than normal

controls. Right-frontal and right-subcortical patients were not significantly impaired in the use of any measured acoustic parameter. However, there was a data trend suggesting that right-frontal and right-subcortical patients used less pitch variation in their speech compared to all the other groups, which was predicted by the right hemisphere and cue-lateralization hypotheses.

Emotion Profiles

Several studies have characterized posed emotions on the basis of different acoustical cue patterns (Ouellette & Baum, 1994; Scherer et al., 1991; Sobin & Alpert, 1999). One of the goals of this study was to evaluate whether the acoustic cues that differentiate posed emotion hold in spontaneous emotional speech. In spontaneous expression, individuals use prosody in conjunction with facial expressions and words to express emotion while in posed prosodic expression actor or subjects are instructed to express emotion through only one channel.

Our results show that “happy” was characterized by mean pitch values that were significantly higher than those samples identified as sad, angry, or unknown. This was true independent of the strength of rater agreement suggesting that this is a core feature in the prosodic channel of expressing happy. Both pitch and volume standard deviation measures were significantly greater for samples identified as happy than for any other category. Finally, samples identified as happy had the highest speech rate of all four categories. Results from mixed-effects analysis were congruent with these findings, revealing that increases in pitch mean, pitch standard deviation and volume standard deviation are significant predictors of happy intensity ratings. Further, increases in speech rate, pitch slope, and pitch regression significantly contributed to the linear model of rated intensity.

“Anger” was also associated with high levels of mean pitch, however unlike happy, mean pitch levels increased significantly as the strength of rater agreement increased. Pitch standard deviation increased significantly from rater agreement strength level two to three for angry samples but not for other categories. This suggests that although increased levels of mean pitch and pitch standard deviation are critical to express happy, for anger, these cues are graded acoustical features that increase as it becomes clearer that one is angry. Also, volume standard deviation was less variable and speech rate were not as fast in samples identified as angry relative to happy samples. Mean pitch, pitch standard deviation, and volume standard deviation were all significant predictors of rated intensity of angry samples, as they were for happy. However, neither pitch slope nor speech rate contributed to the angry statistical model they had for happy and may be a source of acoustic differentiation between the two emotions.

Finally, “Sad” was quite different on almost all acoustical parameters. Sad samples had the lowest mean pitch value of all the categories including samples that were labeled as “unknown”. Further, pitch standard deviation declined as the strength of rater agreement increased and the downward pitch regression slope typical of happy and angry samples was significantly less pronounced in samples identified as sad. Additionally, the volume standard deviation and speech rate also resulted in the lowest mean values for sad samples. Lastly, the number of long pauses in the sample was the best predictor of sad intensity followed by the pitch regression slope in the first half of the sample and mean pitch. Interestingly, most analyses in this study that were significant for the number of long pauses were also significant for the long pause rate, but this was not the case here. This may suggest that there is a cumulative effect of long pauses on the meaning of emotional speech that is lost when looking solely at the rate of pauses.

These findings generally parallel the posed expression and perception literature (Banse & Scherer, 1996; Breitenstein et al., 2001; Pell, 2001; Shapiro & Danly, 1985; Sobin & Alpert, 1999; Van Lancker & Sidtis, 1992) suggesting that studies of posed and spontaneous emotional expression are comparable and do not represent a fundamental methodological shift in the study of emotional production. One point of disagreement in the literature is the speech rate difference between happy and angry. Both Sobin and Alpert (1999) and Breitenstein (2001) characterize angry as having a faster rate of utterance than happy whereas the current results are consistent with Pell (2001) finding that happy vocalizations were faster than angry. Additionally, Sobin and Alpert (1999) found that volume variance was "high" for angry and sad and "moderate" for happy. Results from the current study suggest that volume standard deviation was highest for happy and lowest for sad. The significance of these differences is unclear as it may represent differences due to the use of actors (Breitenstein et al., 2001; Sobin & Alpert, 1999) versus untrained research subjects (current study; Pell, 2001). On the other hand, this difference could be a function of studying different shades of emotion. For instance, Banse and Scherer (1996) noted that discrepancies among studies regarding angry may be attributable to the difference between "cold" anger and "hot" anger. These discrepancies merit further study.

Gender specific emotion profiles

The general prosody literature of speech is unclear about the expected differences in acoustic parameters between men and women (Fitzsimons, Sheahan, and Staunton, 2001), with even less known about gender differences in the prosodic expression of emotion. Results from this study strongly suggest that men and women use pitch and duration cues differently to express emotion. Fitzsimons, Sheahan, and Staunton (2001)

evaluated gender differences in the use of acoustics to express declarative and interrogative intonation. They found that speech rate measures were faster for men than for women and that men increased their speech rate to express an interrogative as opposed to a declarative sentence, and that women did not modify their speech rate. It was also found that men had a significantly greater pitch range and pitch slope. Further, men modified their pitch slope more than women when expressing an interrogative over a declarative sentence.

Results from the current study corroborate the findings of Fitzsimons et al. (2001) and extend them to the expression of emotion. Women, not men, were found to significantly increase their mean pitch from neutral levels to express happy and angry. Similarly, women used gradual changes in pitch regression over the course of speaking (e.g., larger fall of pitch for happy and angry compared to neutral and less fall of pitch for sad compared to neutral). On the other hand, men showed a greater increase in pitch standard deviation to express emotion than women and showed significant increases in pitch slope from neutral levels to express angry. Finally, only men decreased the amount of silence time and long pauses from neutral levels to express angry. These results suggest women primarily rely on the modification of pitch level to express emotion while men rely on tempo changes of pitch and duration cues as well as modulating pitch range. Together with the findings of Fitzsimons et al. (2001), these results strongly suggest that men and women use pitch and duration cues very differently in prosody to communicate emotional messages.

Also of interest, men appeared to increase pitch slope to express the emotion sad, in contrast to women who decreased pitch slope. To express sad, men decreased the amount of silence time from neutral levels and women increased silence times and rate of

long pauses from neutral levels. Rating data suggested that when intending to express sad, women were more accurate than men. It is possible that men and women use pitch slope differently to express the emotion sad, but it is also possible that they were expressing qualitatively different sad. For instance, men may have been more “upset” than sad in which case an increase in pitch slope makes perfect sense. Systematic examination of the content of the stories would need to be conducted to shed light on this possibility and may warrant study in the future.

Women sound happy and men sound angry

Multiple analyses suggest that when speaking neutrally women tend to sound happy and men tend to sound angry. The identification rating data showed elevated agreement for samples judged happy for women and angry for men. The fact that the identification rating difference-score analysis was not significant suggest that this difference between men and women is not a function of women expressing the emotion of happy more effectively than men and vice versa for angry, but that the baseline speech of women and men gravitates toward sounding happy and angry respectively. This interpretation is bolstered by the finding that a significantly greater proportion of *all* samples produced by women were agreed to sound happy while a greater proportion of samples produced by men were agreed to sound angry. The data pattern from the confidence ratings also provided corroborating evidence to this interpretation. Results revealed that raters were most confident when rating angry samples from normal control men and rating happy samples from women from all groups. Further, results indicate that normal control men expressed the emotion angry better than men in other groups or women. Happy samples from normal control women generated greater rater confidence compared to sad and angry samples.

The existence of this effect is supported by acoustical cue production differences between the sexes. Women significantly increased their pitch level from neutral levels to express happy and men did not. As described in the previous section, increased pitch is a stable feature in the expression of happy versus other emotions (Banse & Scherer, 1996; Pell, 2001; Sobin & Alpert, 1999; Breitenstein et al., 2001; Shapiro & Danly, 1985; Van Lancker & Sidtis, 1992). The physical fact that women's voices are higher in pitch than men's gives credence to the possibility that women may sound happier in general. The tendency of men to have a faster pitch slope, increased pitch standard deviation, and less pausing than women is consistent with the expression of angry emotions (Banse & Scherer, 1996; Breitenstein et al., 2001; Pell, 2001; Sobin & Alpert, 1999; Van Lancker & Sidtis, 1992).

A possible confound is that a rater gender bias was active since all the judges in this study were female. Hypothetically, it is possible that male judges would have perceived the samples differently. Rating data showed that samples produced by females were judged to sound like the intended emotion of the subject more often than samples produced by male subjects. Although suggestive of a gender/perception bias, this fact does not fully explain the observed effect. Further, Pell (2002) conducted a study on nonverbal (prosody and face) emotion using both male and female judges. There were no judge x gender biases reported for either voice or face expression. Furthermore, the gender specific acoustical profiles, which are not subject to judge bias, conform to what is predicted if women tended to sound happy and men tended to sound angry. This expression difference between gender may be active during natural spontaneous expression and therefore go undetected in prosodic expression research relying on posed expression. The paucity of studies using spontaneous emotional prosody makes it

reasonable to think that this gender effect is present yet has gone unreported in the literature due to methodological issues.

Perhaps there is an advantage for women to sound happy and men angry. It is commonly accepted that women's pitch is significantly higher than men's (Fitzsimons et al., 2001). Current theories of explanation for this difference include structural and physiological differences in the larynx and pharynx, social influences including group/class membership and differentiating personal identity, and those due to effects of the interlocutor (Lewis, 2002). However, theorizing from a neurobiological and neuropsychological perspective, an alternate perspective to gender differences of pitch and prosody is to examine potential advantages to each gender based on diverging sexual reproduction roles.

Among all mammals that have been closely studied, the female brain is more prepared than the typical male brain to care for infants (Panksepp, 1998). Cells of dorsal preoptic area control nurturance behaviors in men and women, but there are more cells in women due to the inductive effect of estrogen (Panksepp, 1998). The higher pitch found in women may reflect an advantage in producing infant-directed speech is characterized by high pitch and exaggerated pitch contours – a finding that has been shown to hold across languages and cultures (Trainor & Desjardins, 2002). A recent paper by Kaplan, Bachorowski, Smoski, et al. (2002) showed that depressed mothers used less exaggerated prosody compared to nondepressed mothers, resulting in less infant response. Possessing a higher pitched voice should be an advantage for producing infant-directed speech and to sound happy under even neutral conditions. Another possibility is that there is a systematic advantage to sound happy under normal conditions while nurturing infants or interacting in female social structures where women often share child care

responsibilities. Of course, sounding happy under neutral conditions could simply be a byproduct of having a higher pitched voice – a higher pitch being the primary characteristic of expressing happy – and there is no systematic advantage to sounding happy per se.

Male evolutionary history, on the other hand, revolves around the defending of resources and territory and the pursuit of mates (Ellis, 2001). Males are generally more aggressive than females due to generally higher levels of testosterone, which mediate aggression in both men and women (Carlson, 2001) and because of fetal organization and adolescent neural activation effects of testosterone (Panksepp, 1998). It may be that male voice characteristics lend themselves to expressing anger or aggression easily, which has a carryover effect into neutral settings, as the tendency of men to modulate pitch slope, range and speech tempo are key characteristics to the expression of angry. It is conceivable that in men's quest for resources and mates, it would be advantageous to have an aggressive tone even in non-aggressive situations.

The interpretation that it is evolutionary advantageous for women and men to use happy and angry prosody in neutral speech respectively is not exclusionary of existing theories for gender pitch differences. Furthermore, other neurobiological interpretations are certainly conceivable as well. In any case, both rater and acoustic data from the current study suggest that women tend to sound happy and men sound angry, consistent with evolutionarily determined sex roles. This effect is worthy of additional study and requires replication using both male and female judges. If this effect is robust, further study of advantages to emotional prosody in neutral speech could inform theoretical explanations for vocal differences in men and women.

Emotional communication hypotheses

The identification rating data provides the best foundation to evaluate the emotional communication hypotheses. Samples produced by both RBDs and LBDs had a higher agreement rate for sad but lower rate for happy and angry samples relative to NCs. A plausible explanation of this finding is that BD subjects in this study were generally perceived to sound depressed, or produced output that lacked emotional prosody and sounded "flat". Follow up analyses revealed that a significantly higher proportion of samples produced by BDs were rated as sad compared to NCs. Further, the intensity ratings did not suggest that BDs expressed the emotion sad more intensely than NC subjects. Taken together, these data support the interpretation that raters interpreted BD patients' prosody as depressed resulting in more samples being judged as sounding sad.

A further breakdown of the identification rating data by lesion site revealed an important and statistically significant pattern of data suggesting that left-frontal patients were the most impaired group followed by right-frontal and right-subcortical patients. With respect to evaluating subcortical hypotheses of prosody, it is important to note that left subcortical patients did not have elevated sad agreement rates. Lastly, the identification rating data also suggested that both left and right temporal patients did not sound normal. Therefore, these results specifically suggest that the performance of left-frontal, right-frontal, and right-subcortical patients account for the elevated sad identification ratings of the BD participants.

Inspecting the pattern of differences in acoustic parameters of the samples provides some explanation to the identification rating data pattern. Left-frontal patients had difficulty with duration cues, having significantly more silence time and a greater number of long pauses in their samples. In general, LBD patients spoke at a significantly slower rate than normal controls. Speech produced by neurologically intact depressed

patients typically contains significantly more pause time and slower utterance duration than speech produced by controls (Alpert, Pouget, & Silva, 2001; Garcia Toro, Talavera, Saiz Ruiz, & Gonzalez, 2000). Current results from mixed-effects analysis revealed that the number of long pauses in a sample was a significant predictor of rated intensity of sad suggesting that this acoustical feature signified sadness to our judges. Together, these results suggest that spontaneous speech produced by left-frontal patients is similar to those with depression. All patients from the current study were screened for depression using the Beck Depression Inventory, the same scale used in Alpert et al. (2001), and were not clinically depressed at time of testing. Therefore, left-frontal patients had speech timing difficulties similar to depressed patients, despite lacking clinical symptoms of depression. This is also an interesting finding in light of findings suggesting that patients with damage to left frontal areas are more depressed than patients with damage to any other part of the brain (Robinson & Manes, 2000). How these findings may be related remains unclear.

The notion that the increased pause rate observed in left frontal patients is speechrelated is consistent with Ross et al. (1997) finding that LBDs performed as poorly as RBDs in language demanding prosody tasks, however when the language demands of the task were reduced, LBDs showed marked improvement in performance while RBDs did not. It is also possible that this finding suggests that the left frontal patients in this study show signs of mild apraxia or dysarthria, as this could not be categorically eliminated as a possible explanation. However, broadly speaking these findings do support the cue-lateralization hypothesis (Schirmer et al., 2001; Van Lancker & Sidtis, 1992) which posits that the left hemisphere is predominantly responsible for the production of duration cues (e.g., pauses) in prosody.

Emotionally “flat” prosody is characterized by reductions in pitch measures (Pell, 1999b). Pell (1999b) identified RBD subjects that were rated as emotionally “flat” and compared them to other subjects. It was found that “flat” subjects had reduced mean pitch levels accounting for being rated emotionally blunted. Ross et al., (1986) states that loss of affective aspects of prosody in right frontal patients is attributable to reduction in dynamic range of pitch while Shapiro and Danly (1985) found that right-frontal and right-central patients had reduced pitch variation and that right posterior and left posterior lesion patients could use pitch variation to express happy and sad. Additionally, Pell (1999a) found that RBD patients used less pitch variation than other groups.

There was not a significant lesion site effect ($p = 0.17$) of pitch standard deviation in the current study. However, the pattern of results showed that the right-frontal and right-subcortical patients had less pitch variation in their speech relative to all the other groups. Had this effect been significant, not only would this finding explain the identification rating effect, but would also strongly support the cue-lateralization hypothesis in conjunction with the finding that left-frontal patients have difficulty with duration cues and the right hemisphere hypothesis. The small number of subjects in each group (i.e., average of 6 subjects per BD group), and the absence of any other acoustical correlate to explain deficits in right frontal and subcortical patients, leaves open the possibility of a type II error.

Ross (1997) concluded that their data supported the supposition that affective prosody is strongly lateralized to the right hemisphere (e.g., right hemisphere hypothesis) after finding that LBD performance on prosody tasks improved when language demands were reduced while RBD performance did not. On that basis, the current results could suggest support for the right-hemisphere hypothesis since deficits observed in left-frontal

patients appear speech related. Furthermore, left-subcortical patients did not have emotional prosody deficits whereas both right frontal and right-subcortical patients did, suggesting a more significant role in emotional prosody for the right hemisphere. Studies have suggested that human raters are more accurate than acoustical analysis at detecting emotional prosody (Rabinov et al., 1995; Sobin & Alpert, 1999) justifying making conclusions on the basis of the identification rating findings alone. However, given the connection between emotional prosody and pitch modulation (Pell, 1999b; Shapiro & Danly, 1985) only had the analysis of pitch standard deviation effect reached significance, strong support for the right hemisphere hypothesis and cue-lateralization hypotheses could be given.

Limitations

Although there are studies that suggest that women are better at detecting emotion than men, differences among the gender groups using rating data are difficult to interpret without the inclusion of male raters in the study. Of the 30 individuals that replied to recruitment advertisements to participate in this study, only 2 were men, making it impossible to include equal numbers of both genders as raters. Acoustical data from the current study and other studies using both male and female raters (Pell, 2002; Scherer et al., 1991) suggest the absence of a rater gender bias, however, had male raters been included this possibility might have been categorically eliminated.

Acoustical analysis was not intended when the original methodological conceptualization for the collection of the spontaneous speech data was created. Unfortunately, recording variations between subjects, which was acceptable for human ratings, did not allow the examination of mean volume differences between samples. Other studies suggest that volume is often a cue used to differentiate different kinds of

emotional expression. Now that acoustical analysis procedures are being used in this laboratory, future studies will include methodology standards that ensure that volume data can be analyzed.

Finally, one difficulty of any lesion study is finding enough subjects with lesions in specific locations. There is enough literature available that suggests that broad left and right lesions are not specific enough to be able to make clear statements about lateralization of function. Although all available frontal and subcortical patients were included, many of the analyses based on specific lesion location were not significant despite predictable and supportable patterns in the data. Future data collected could be added to these analyses to ascertain whether trends observed with acoustical data are true effects.

Future Directions

Future studies might benefit from using subjects closer to stroke onset. Baum and Pell (1999) suggest that months post onset may prove to be a critical variable in the emergence of dysprosody. Several of the reviewed studies finding differences examined patients in the relatively acute stage, with the longest time since stroke onset being three months (Edmondson, Chan, Seibert, & Ross, 1987; Gandour et al., 1995; Hughes, Chan, & Su, 1983). Many of the patients studied in this experiment were well beyond three months-post-onset of stroke and may have recovered a good deal of prosodic functioning. Further, patient deficits were required to be relatively mild in order to pass cognitive screening measures. Together, this could suggest that any deficits our patients exhibit might be too mild to detect statistically significant differences.

Another point worth exploring is the methodology used to elicit emotional responses from the patients. Despite patients' high level of self-rated emotional intensity

while re-telling emotional stories from their past, experimenter observations and rating data from the current study suggest that patients may have had a low level of emotional involvement. Future work might benefit from using other kinds of elicitation procedures that generate more genuine emotional response from subjects. Analyzing other emotions (e.g., fear) could also reveal important information about the neural mechanisms of prosody production.

Lastly, future studies could benefit from other types of analyses. For example, it could be beneficial to understand the relationship between individual rater confidence and the number of raters that agree that a prosody sample sounds like a particular emotion. Another beneficial analysis to conduct is to evaluate the identification rating effect using the long pause rate as a covariate. If the identification rating effect seen for left frontal patients became less pronounced or disappeared entirely while the effects seen in right-hemisphere patients maintained itself, stronger support for the right hemisphere hypothesis could be given. Lastly, non-linear mixed effects models examining the predictability of rater identification judgments from sample acoustical parameters would shed light on which specific acoustics raters listen to when making judgments about different emotions.

Conclusions

Significant relationships between rater judgments and sample acoustics suggest that emotional profiles based on posed emotional expression are valid in spontaneous emotional expression as well. This is meaningful as posed expression studies are controlled studies intended to generalize to real situations. The results of this study suggest that this is generally appropriate. One area where this may not be appropriate is the comparison of men and women. Results strongly suggest that men and women use

pitch and duration cues differently to express emotion. Given that these findings have been found in several studies, clinical prosody retraining programs should take these differences into account.

Furthermore, results suggest that men's voice characteristics may contribute to sounding somewhat angry under even neutral conditions; similarly women's characteristics contribute to their sounding happy. The fact that this pattern has not been found in posed studies could mean that this effect does not occur under posed conditions or that language cues that are included in the rating of posed prosody mask such an effect. Further study using spontaneous emotional speech should be conducted to ascertain whether this is a robust effect. Taken together, enough differences were found between the gender groups that future studies should be wary of using men and women as identical units of study.

Finally, these results provided differential support for the emotional communicational hypotheses. First, the valence and motoric direction hypotheses are clearly not supported since there is no evidence that any BD group produced prosody normally for one emotion and abnormally for another. Patient groups showing prosodic abnormalities appeared to have them regardless of the emotion being expressed, contrary to the predictions of both hypotheses. The right hemisphere hypothesis was partially supported by the identification ratings suggesting that left-frontal, right-frontal, and right-subcortical patients had prosody deficits. The facts that deficits of left-frontal patients were attributable to slowed speech and increases in pauses and left-subcortical patients did not show deficits suggest that the right hemisphere may be more involved in emotional prosody production than the left hemisphere. The cue-lateralization hypothesis is partially supported since left-frontal patients showed duration cue difficulties and,

although not significant, there was a trend for right-frontal and right-subcortical patients to have difficulty modulating pitch range. Lastly, these data also supported findings that the fronto-striatal circuitry, not just subcortical structures (Cancelliere & Kertesz, 1990), are critical for prosody production (Breitenstein et al., 1998), with the caveat that the right hemisphere may be more important than the left hemisphere.

Appendix A.

Pilot procedures

Monologue sampling pilot procedures

Overview. The monologue sampling procedure was piloted to ensure that reliable reproducible samples could be extracted from the monologues using this procedure. This process involved two phases: 1) simulating the proposed procedure by having two experimenters make independent judgments on 12 (i.e., 3 happy, 3 sad, 3 angry, and 3 nonemotional supermarket) sample monologues produced by brain-damaged and neurologically intact subjects (i.e., 3 RBD, 3 LBD, and 3 NCs) and tabulating the agreement rates; and 2) 3 Ph.D. students learned the criteria and applied it to the same 12 monologues. Agreement rates between each student and the experimenters were calculated to provide evidence that outside persons can learn and apply the monologue categorization criteria and is therefore objective.

Sampling procedure review. A detailed explanation of the proposed procedure can be reviewed on page 108. To summarize these procedures, experimenters performed a sentence by sentence analysis and categorized each sentence as “context” (not rated), “play-by-play” (rated), “reflection” (rated), or “closing” (not rated) sentences based on the target event that is stated to elicit the target emotional experience and the temporal relationship the information being categorized has to this target event. Based on independent implementation of these procedures on the sample monologues, interrater reliability was calculated via complete agreement.

Reliability results. The reliability results were very good for both phases of the pilot. Phase 1 results show that experimenters LC and MS had an average agreement rate of 83.1% on 12 monologues, leaving 16.9% of the monologues in disagreement (see table 6 below).

Table 6: Agreement rates by subject

Emotion	Words	Agreement			Disagreement type					
		Agree	Disagree	%	P*R	P*R %	N*P	N*P %	N*R	N*R %
Happy	248	187	61	75%	34	14%	27	11%	0	0%
Angry	432	318	114	74%	0	0%		0%	114	26%
Sad	448	375	73	84%	18	4%	20	4%	35	8%
Happy	405	333	72	82%	0	0%		0%	72	18%
Angry	368	207	161	56%	102	28%	17	5%	41	12%
Sad	107	107	0	100%	0	0%		0%	0	0%
Happy	195	175	20	90%	20	10%		0%	0	0%
Angry	289	262	27	91%	0	0%	11	4%	16	6%
Sad	206	152	54	74%	44	21%	2	1%	8	4%
Market	389	280	109	72%	0	0%	20	7%	83	21%
Market	215	215	0	100%	0	0%		0%	0	0%
Market	165	165	0	100%	0	0%		0%	0	0%
83.1%					6.4%		2.6%		7.8%	

P*R = same segment categorized as "play-by-play" by one experimenter and "reflection" by the other;
N*P = same segment categorized as "not rated" (i.e., context, closing, excluded due to cue) by one experimenter and "play-by-play" by the other rater; N*R = same segment categorized as "not rated" (i.e., context, closing, excluded due to cue) by one experimenter and "reflection" by the other rater.

For each monologue, agreement was calculated by taking the total number of words and dividing it by the number of words that the experimenters categorized identically. This calculation yielded a percent agreement for each monologue, which was averaged with the 11 other monologue agreement rates to generate an overall agreement of 83.1%.

Breaking down overall disagreement, 6.4% of the disagreement was attributable to the same segment being categorized as "play-by-play" by one experimenter and "reflection" by the other. In this case, both experimenters designated the segment as a category that should be rated for the target emotion and therefore this pilot demonstrated an 89.5% agreement rate on what to include for rating and what to exclude from rating. The remaining disagreement is attributable to the same segment being categorized as "not rated" (i.e., context, closing, or excluded due to cue) by one experimenter and "play-by-play" by the other rater (2.6%) or categorized as "not rated" by one experimenter and "reflection" by the other (7.8%). This discrepancy in error frequency is consistent with

the fact that "play-by-play" segments are less ambiguous to categorize than "reflection" segments. Furthermore, similar rates of agreement were consistent across subject group and emotion (see table 7 below).

Table 7: Agreement rates by emotion and brain-damage classification

Emotion	Words	Agreement			Disagreement type					
		Agree	Disagree	%	P*R	P*R %	N*P	N*P %	N*R	N*R %
Happy	248	187	61	75%	34	14%	27	11%	0	0%
Happy	405	333	72	82%	0	0%	0	0%	72	18%
Happy	195	175	20	90%	20	10%	0	0%	0	0%
				82.3%		8%		3.7%		6%
Angry	432	318	114	74%	0	0%	0	0%	114	26%
Angry	368	207	161	56%	102	28%	17	5%	41	12%
Angry	289	262	27	91%	0	0%	11	4%	16	6%
				73.6%		9.3%		3%		14.1%
Sad	107	107	0	100%	0	0%	0	0%	0	0%
Sad	448	375	73	84%	18	4%	20	4%	35	8%
Sad	206	152	54	74%	44	21%	2	1%	8	4%
				86%		8.3%		1.7%		4%
Market	389	280	109	72%	0	0%	20	7%	83	21%
Market	215	215	0	100%	0	0%	0	0%	0	0%
Market	165	165	0	100%	0	0%	0	0%	0	0%
				90.6%		0%		2.4%		7%
NC	248	187	61	75%	34	14%	27	11%	0	0%
NC	432	318	114	74%	0	0%	0	0%	114	26%
NC	448	375	73	84%	18	4%	20	4%	35	8%
NC	389	280	109	72%	0	0%	20	7%	83	21%
				76.3%		4.5%		5.5%		13.7%
LBD	195	175	20	90%	20	10%	0	0%	0	0%
LBD	289	262	27	91%	0	0%	11	4%	16	6%
LBD	206	152	54	74%	44	21%	2	1%	8	4%
LBD	165	165	0	100%	0	0%	0	0%	0	0%
				88.6%		7.7%		1.2%		2.5%
RBD	405	333	72	82%	0	0%	0	0%	72	18%
RBD	368	207	161	56%	102	28%	17	5%	41	12%
RBD	107	107	0	100%	0	0%	0	0%	0	0%
RBD	215	215	0	100%	0	0%	0	0%	0	0%
				84.5%		7%		1.2%		7.3%

NC = normal control subject; LBD = left brain-damaged subject; RBD = right brain-damaged subject; P*R = same segment categorized as "play-by-play" by one experimenter and "reflection" by the other; N*P = same segment categorized as "not rated" (i.e., context, closing, excluded due to cue) by one experimenter and "play-by-play" by the other rater; N*R = same segment categorized as "not rated" (i.e., context, closing, excluded due to cue) by one experimenter and "reflection" by the other rater.

Phase 2 results show that the 3 Ph.D. students reliably applied the criteria in a manner consistent the experimenters (see table 8 below). These results reveal a high level of agreement for what should be rated and what should be excluded (e.g., agreement with Experimenter 1: 71%, 74%, 77%, and agreement with Experimenter 2: 76%, 76%, 77%). The only training the students received was reading Appendix A, which describes how to categorize sentences based on the sentence “types” and criteria.

Table 8: Experimenters (MS/LC) agreement rates with students (JB, SP, AR)

Experimenter 1 (MS) agreement rates with students (JB, SP, AR)										
Emotion	Words	JB A	JB NA	JB %	SP A	SP NA	SP %	AR	AR NA	AR %
Happy	248	201	47	81%	225	23	91%	171	77	69%
Angry	432	391	41	91%	287	145	66%	304	128	70%
Sad	448	299	149	67%	214	234	48%	374	74	83%
Happy	405	224	181	55%	268	137	66%	268	137	66%
Angry	368	345	23	94%	128	240	35%	283	85	77%
Sad	107	94	13	88%	102	5	95%	102	5	95%
Happy	195	195	0	100%	195	0	100%	195	0	100%
Angry	289	220	69	76%	218	71	75%	229	60	79%
Sad	206	206	0	100%	154	52	75%	148	58	72%
Market	389	234	155	60%	219	170	56%	207	182	53%
Market	215	73	142	34%	178	37	83%	175	40	81%
Market	165	120	45	73%	99	66	60%	104	61	63%
							77%			76%
Experimenter 2 (LC) agreement rates with students (JB, SP, AR)										
Emotion	Words	JB A	JB NA	JB %	SP A	SP NA	SP %	AR	AR NA	AR %
Happy	248	174	74	70%	198	50	80%	144	104	58%
Angry	432	359	73	83%	314	118	73%	418	14	97%
Sad	448	354	94	79%	266	182	59%	360	88	80%
Happy	405	296	109	73%	339	66	84%	340	65	84%
Angry	368	291	77	79%	186	182	51%	248	120	67%
Sad	107	94	13	88%	102	5	95%	102	5	95%
Happy	195	195	0	100%	195	0	100%	195	0	100%
Angry	289	193	96	67%	245	44	85%	176	113	61%
Sad	206	196	10	95%	143	63	69%	143	63	69%
Market	389	275	114	71%	305	84	78%	281	108	72%
Market	215	73	142	34%	178	37	83%	175	40	81%
Market	165	120	45	73%	99	66	60%	104	61	63%
							76%			77%

MS and LC are the initials of the experimenters; JB, SP, AR are the initials of the students; A = words in agreement; NA = words in disagreement.

Rating emotional accuracy piloting

Overview. To address questions regarding acoustical correlates of specific emotions and rater acoustical property relationships, it is important that some of the filtered monologue samples will be accurately identified by judges. Since one of the proposed goals of this experiment is to establish relationships between rater judgments of monologue samples and acoustical parameters, it is important to establish that raters can agree on what emotion the output sounds like. Without agreement on a subset of samples, these relationships will be impossible to explore. Two experimenters categorized 3 angry, 3 happy, and 3 sad monologues (85.6% agreement), produced by 5 men and 4 women from written transcripts of normal control subjects from the larger study using the procedures outlined above to collect a sample of rating data to ensure the likelihood of interpretable results. Only normal subjects were used for this pilot because it is possible that brain-damaged subjects will not be able to produce emotional prosody accurately and therefore it is uninformative to the purposes of this pilot to sample whether judges can accurately identify brain-damaged samples. Selected speech samples were cut in 15 to 25 second samples from the entire monologue. Based on previous work, this sample time range allows raters enough time to judge the sample and also allows longer speech samples to be broken into parts (Speilman, 2002). For example, a 30-second speech sample was cut into two 15-second speech samples. In total, 8 angry segments (5 "play-by-play", 3 "reflection"), 9 happy segments (6 "play-by-play", 3 "reflection"), and 10 sad segments (2 "play-by-play", 8 "reflection") were presented to judges for ratings.

Digital processing of speech samples. Speech samples were recorded from audio tape to computer through a sound blaster soundcard at 22,000 Hz. The samples were cut into 15 to 25 second samples using Creative Wave Studio 4.08 and then Praat 6.08

speech analysis software was used to low-pass filter male voices at 425 Hz and female voices at 575 Hz, to eliminate 95% (Speilman, 2002) of the lexical content of speech while leaving the prosodic intonations essentially intact. Finally, each sample is equalized for loudness using a script run in Matlab 6.01 which attenuates and/or amplifies each signal to reach an average of 70 dB. This is necessary because loudness does affect perceived intensity, and the samples are rather variable in loudness due to testing conditions. This equalization removes loudness differences as a source of variation.

Rating procedure. For the purposes of this pilot, raters were not formally trained and were asked to listen to filtered speech samples and make 3 judgments including: how emotionally intense they thought the sample was; whether it sounded sad, happy, or angry (neutral was not an option); and how confident they were in their emotional accuracy rating. The raters used a 7-point Likert intensity scale (Canino et al., 1999) and a 5-point Likert confidence (Borod et al., submitted for publication; Canino, 2001) scale used in prior research in Dr. Borod's laboratory, which are proposed for the experimental rating session and described in detail below. Speech samples were presented in random order up to two times.

Data coding. The monologue categorization process is a method of identifying sentences from a monologue to be rated that are the most likely to express the target emotion prosodically, and therefore produce a suitable speech sample of the target emotion. As such, it is not only likely, but expected that some of the rated segments will be neutral in tone or reflect a different emotion than the target emotion entirely. As such, the segments derived from a subject's monologues were ranked based on which segment achieved the highest agreement as reflecting the target emotion. The rationale is that the sample that has the highest agreement rate for the target emotion is the best sample of the

target emotion for each monologue and therefore should be used as the primary representative of that monologue. This process eliminates the potential washing out of a good target emotion sample that could occur by treating all samples identically regardless of what emotion they were rated as.

For example, Subject 120 produced a happy monologue of approximately 2 minutes and 45 seconds. Based on the monologue categorization, approximately 20 seconds of the monologue were "reflection" and approximately 45 seconds were "play-by-play". The 45 second segment was then broken into three 15 second segments to present to the raters. The 20 second "reflection" segment does not need to be rank ordered since there is only 1 segment presented to subjects, while the three 15 second segments do need to be ranked to identify which of the 3 is the best sample of the target emotion, based on the raters' judgments.

Table 9: Ranking example

Subject	Category	Sad	Angry	Happy	Agree	Rank
120h	play-by-play		0.33	0.67	0.5	1
120h	play-by-play	0.33	0.67		0.5	x
120h	play-by-play	0.67	0.33		0.5	x

The table above highlights the ranking process. In this case, the first segment was perceived as angry by one rater and happy by two raters. The second segment was perceived as sad by one rater and angry by two raters, while the third segment was perceived as sad by two raters and angry by the last rater. As such, the first segment is ranked 1 and is used for calculating accuracy rates. For the purposes of this pilot, no analyses were performed on unranked samples.

Rater reliability Results. Independent of the accuracy of their judgments, raters agreed on the emotion angry segments sounded like 75.89%, happy segments 59.22%, and sad segments 66.6% with chance being at 33%. More specifically, raters agreed on

angry “reflection” segments 100% and angry “play-by-play” segments 63.83%; happy “reflection” segments 61% and happy “play-by-play” segments 58.33%; and sad “reflection” segments 70.75% and sad “play-by-play” segments 50%. Rater confidence does appear to be related to overall agreement as well as the perceived sample intensity (see table 10 below).

Table 10: Relationship between rater confidence and agreement/intensity

Average Rater Confidence	Agreement Percentage	Average Perceived Intensity
1	50.00%	2.77
2	65.31%	3.23
3	71.18%	3.66
4	100.00%	5.00

Rater accuracy Results. The accuracy rating was evaluated using the top-ranked data only revealing that raters correctly identified angry and sad samples at a rate of 61.11% and 66.67% respectively while at a rate of only 22.22% for happy samples. Perceived intensity did not appear to be related to the correctness of rater’s accuracy judgments. For a breakdown of rater performance for each emotion and category, see the table 11 below:

Table 11: Percent agreement by emotion

Target Emotion	Category	Percent correctly identified
Angry	“Reflection”	66.67%
Angry	“Play-by-play”	55.56%
Happy	“Reflection”	11.11%
Happy	“Play-by-play”	33.33%
Sad	“Reflection”	86.67%
Sad	“Play-by-play”	16.67%

Pilot Discussion. These results suggest that raters should achieve an acceptable level of inter-rater agreement and that emotional segments can be correctly identified using the above procedures. Of potential interest is the difference between the “reflection” and “play-by-play” percentages when comparing angry to sad which possibly

suggests that prosody is used only during reflective statements for the emotion sad but is used for both statements when expressing angry.

However, a concern is that raters correctly identified happy segments at only chance or below chance levels. A potential explanation for this is that subjects are not asked to “sound” happy during the monologue since it is spontaneous rather than posed. It is possible that subjects rely on different channels other than prosody to convey happiness spontaneously. In support of this explanation is the fact that raters continued to achieve similar agreement levels for happy segments relative to angry and sad segments despite the decreased accuracy of their assessment.

Further, the experimenters evaluated the unfiltered samples to assess whether the segments sounded happy, in order to determine if the failure of the raters to identify happy accurately suggested that they needed more training. The filtered samples did not sound happy to either experimenter, and an appraisal of the unfiltered samples revealed very subtle happy sounding prosody. In contrast, the subject’s verbal content was vividly happy, however he only expressed a slight up tempo stressing of certain words. Several happy samples were similarly evaluated with the same basic finding. This subtly is apparently filtered out along with the verbal content of the samples and therefore probably explains why the happy samples are not identified well. There were also no other portions of the monologues that sounded happier than the samples we collected and so it does not appear that monologue categorization procedure misses the happy sounding samples.

Lastly, angry “play-by-play” and “reflection” samples were accurately identified similarly, while sad “reflection” samples were accurately identified at a much greater rate than sad “play-by-play” samples. Together with the happy results, this pilot suggests that

the use of prosody in spontaneous emotional communication may be different for each of these emotions. This pattern of results will be predicted for the larger study, and if confirmed, will inform models of the practical use of prosody in emotional expression.

Appendix B

Instructions for selecting discourse from monologues

Subjects were instructed to talk about emotional experiences involving each emotion (i.e., happy, sad, angry) one at a time. They were to try to remember and actually relive the last time that they felt the target emotion, to put themselves back into the situation, and to try to re-experience it with as much real feeling and intensity as when it actually occurred, and were instructed to speak for at least two minutes but not more than four minutes. Their spontaneous production is the target for this analysis. This document addresses the issue of deciding what parts of the monologue should be examined since they are filled with content other than the event(s) that is supposed to elicit the target emotion.

For example, for the target emotion anger, one subject explained that his brother had died recently (which sounded sad of course) and this was why he had to sell this piece of property. What made him angry was how the realtor had treated him, which he went on to explain. So, if we rate the entire monologue, we will be mixing the target emotion, other emotions, and neutral speech together while trying to evaluate only the target emotion. Ideally, we would like to limit the evaluation to discourse that should reflect the target emotion and eliminate the rest of the discourse from analysis in order to restrict data variability and increase the likelihood of finding significant effects, should they exist.

What follows are criteria that were developed for objectively pulling out sentences of the monologue that are related to the event that elicited the target emotion, and therefore need to be rated. Overview of monologue breakdown criteria:

Categories

The monologue sampling procedure categorizes the discourse into four basic parts based on the story structure, which parenthetically is not always sequentially ordered in the monologue, including:

- e) Context: Information leading up to the emotion eliciting event or context to the event that furthers the understanding of the story.
- f) Play-by-play: The actual event or series of events that produced the target emotional experience.
- g) Reflection: Discourse that expresses what the experience meant to the subject, why it made them feel the way they did, or how the specific experience represents their feeling of the target emotion in a larger sense.
- h) Closing: What happened after the event and/or summing up the story, independent of why the subject felt the way they did.

Discourse that conveys the “play-by-play” of the event that produced the target emotion, and also any discourse that explains or expresses the impact the event had on them, or that elaborates on, or implies, how or why the event was meaningful to them (“reflection”) will be included for analysis. These two types of segments are included because they are the most likely to include emotional content, should the subject be emotional at all during the monologue. On the other hand, discourse that provides “context” to the event that produced the target emotion will be excluded from rating. Further, discourse related to what happened after the event and summing up of the story (“closing”) will also be excluded, provided the discourse is **not** related to how the subject

felt or what the event meant to them, in which case it is "reflection" and will be included. Several criteria have been created to help categorize the monologue into these four components and are explained in detail below.

Criterion Number 1: Temporal. All the events that are conveyed in the monologues are from the past. However, since the subjects are supposed to be recalling a specific point in time that they felt a particular emotion, the story usually temporally revolves around the emotional event(s). Generally, discourse that should be categorized as "context" occurs temporally before the event that produced the target emotion. The "play-by-play" occurs in the present tense of the story and the "closing" occurs in the future or includes statements unrelated to the target event. "Reflection" statements are generally not time specific, and it is therefore crucial to identify the target event to be able to assess whether the reflective comments refer to the target event or to something else. For example, from a subject with target emotion happiness:

Well the most recent time was last uh, Sunday, no, last Saturday. I uh, usually try to see a play every once in a while, and one of my best places uh, I enjoy very much is the XXX School of Drama. And they have uh, maybe 5, 6, 7, 8 shows, every uh, every semester, which is performed by the uh, graduate students of the school of down. It's the XXX School of the Arts, on XXX near the XXX. Well, last Saturday they had a play called XXX, and I never enjoyed myself more than to see the...there were 15 people and it was a story about uh, a uh, XXX Church. I mean the scenery wasn't...I mean they didn't have to have very good scenery, but they had a piano and they were playing and singing and the people were very happy to...and they were s...s...they...they rev...revered snakes. Uh, and like I said I felt very good after seeing that performance. (C - Who did you see it with? A need a little bit more.) I went by...I usually go by myself. Uh, there's about 200 seats, most of them are students, but a lot of senior ci...citizens go to see these performances. (C - I need a little bit more. Describe the play a little bit more) Well, it...it uh, the main...it concer...the main character's concern was uh, a man and his wife, and he...she beat 'em, and she went back to 'em, and...and then they uh, had all the other people uh, praising God, and they were very, you know, very emotional. (C - What did you do after?) Oh, then after, I uh, took the train and went home.

As you can see, the selection begins when the event that produced happiness begins, even though to understand the story you would like to know the information that precedes it. You will notice that the 'text' selection explains the details of the event while the 'text' selection explains why the event made them feel happy. We are not interested in the story per se; we are simply using the story structure as a guide to select the speech directly related to the task (generating spontaneous emotional dialogue). Here is another example, but for sadness:

Last night on TV there was a story called XXX Romeo and Juliet. It dealt with this, I guess two...two. A boy and a girl of different nationalities, uh, fell in love with each other. And, uh, they, the parents, knew that they were really in love despite the fact that they were conflicting nationalities. And, uh, they lived together in XXX and in the midst of all this misery and death and... The culminating episode, I'll come back to other stuff, but it's when they tried to get out of XXX and as they were crossing a plaza, first a bullet killed him and she was fatally wounded and as she crawled over to touch him she died. This is true. This is not uh...this really happened because you could see the pictures of the bodies on the plaza. Very moving, disturbing thing. **The, uh, whole business of XXX is a mess, the, uh, way they people have been suffering, the way they've been butchered, the way we are bollicksing up the, uh, arrangements there. It made me feel sad. It made me...it was, you know, these two lovely young people. He was a nice healthy looking boy, and she was a very sweet looking girl, and they're...they were really in love and nothing infatuation. They didn't want to get married because, uh, they were afraid what the relatives would say, but it it's an important situation. There's so...so much blood shed in that area. There's so much bloodshed all over the world, my goodness. Uh, in Africa, you name it they...people being killed. It's, uh, distressing.**

Notice here that the details about the eliciting event (in 'text') was split by discourse explaining how the target event (the movie) elicited a broader feeling of the target emotion (the current state of Bosnia) included in 'text'. Once she made those comments she continued to explore the eliciting event further (again in 'text') which is why it is selected, followed by another round of reflection (in 'text'). It should also be noted that you do not have to have both 'text' and 'text' segments within the monologue. Look at these two examples:

We have neighbors, the XXX. Mr. and Mrs. XXX. With a name like that they still are XXX and they are knockers-up. In other words, if we make any un...un...undue noise, uh, Mrs. XXX will take a broomstick and knock up. She says her ceilings are all pot-marked with, uh, pot, uh, marks. Quite a while ago uh. I had to repair my bed. My bed is in uh, separate bedroom from my wife's and turning the mattress over...off to get it off the bed. I banged the floor. Now, this is a post-war house, lousy apartment house, and it wasn't put up together so well. Sounds and shocks go through the building very regularly. Seems that the shock that I got from the uh, made...created when I dropped the mattress was apparently transmitted to the ceiling of the Mrs. XXX apartment and it broke a uh, little lamp on the...a shade on the lamp in the bedroom there. He came back and he banged on the door. Naturally when he banged on the door I didn't open the door. And my wife tells me he...he issued a curse. He cursed all my entire family, my wife, myself, and my daughter. I can't see that. He transmitted his tale to the superintendent and superintendent got in touch with the landlord got in touch with us that he's gonna sue us for, uh, repairing the lamp. He hasn't done that yet. And she continues to knock up. I dropped a jar on the floor accidentally yesterday. Fortunately it didn't break, but immediately bang, bang she up...she came up, she banged. Recently I found the key to his mailbox. Now, uh, I knew uh...I didn't know uh, his mailbox...my wife was with me and she realized it was XXX. It was his apartment number. So we gave the key to the porter and then when we saw Mr. XXX, XXX told him about it, he was very appreciative about it, but this didn't stop her from knocking up the next time we made noise.

I would have to say it was at the death of a little girl, XXX uh, XXX. First, it was tragedy to me as a father who has 2 girls, and who've gone to mature in adulthood now. To think of a fa...a mother or a father who would beat a child to death. Then there was the tragedy of hearing the mother say that the child was demon-possessed. We deal with this a lot, of course, in being an elder in my church. One of the tragic things you have to point out to people is...is that despite the fact that the Bible says that demon possession is so, there is no way from any scriptural text that I know of, that you could possibly know when or if a person was demon-possessed. That's the problem. Uh, those of course of my background say, Jesus would obviously know, and maybe the apostles would know, but we have been given no information as to how to identify it. Then the second thing that was so...s...that was really so sad to me, was the whole rain dance with these social workers. I think that we've gotten under the impression that uh, social workers are supposed to be...have God-like characters. I...I say to people now and then, when somebody says, well you know, I say, no. I failed XXX 101 and probably will fail XXX 102. Until you tell me. I don't know. And the sad thing is to see all of these officials, from the governor and the mayor on down, say, well we could have done what...done something. Done what? Parents are supposed to take care of their children. Parents are the one who raise them. The parent was there with the child and the parent basically didn't do what was a responsible thing. Now all we can do is react. **So the whole system, it just sad to see, the whole system. But quite frank...if you know, study the communist system, governments don't make very good parents. We've proven that all...all along. What we've gotta do is make parents good parents. So to me**

it was just so sad to see the whole thing. And of course the death was just tragic. And, yes, the mother's responsible. All the evidence points to her, and probably the stepfather is too. But the sadness is in just the way in which we, a grown intelligent people, treat it.

In order to make these kinds of distinctions, you need to be able to identifying the eliciting event or focus of the story, which introduces the next criterion.

Criterion Number 2: Identifying the Key Event. The key event of the monologue is the situation or activity that the subject reports produced the target emotion. Identifying the key event helps categorize the discourse correctly, particularly for the "reflection" categorization where the discourse must be reflection about this event, while reflection independent of this event is either "context", "closing", or extraneous depending on the temporal criterion. In cases where the key event is ambiguous, the temporal criterion takes priority. Refer to Appendix A for specific examples of how the Temporal criterion facilitates the identification of the Key Event criterion in these cases. For example, an angry monologue:

Time I remember, I got very upset and angry. Well happened about 1979. I...I was working for the XXX Government, with the XXX, overseas. And I was...I became XXX, maybe uh, 2 or 3 months before. *At that time, the agency decided that it didn't want anybody over the age of XXX to be in the foreign service, and so uh, there was a court case. Somebody brought suit, and the original court said that the uh, government can't make anybody uh, leave their job because of age. But it went to the XXX Court. And the XXX Court said government can do whatever they want. Uh, they...they can make the rules and regulations. And I had just comin' back from home leave and I figured I'm gonna work coupla more years. But then when they had this decision, they gave us 2 weeks. They said in 2 weeks you have to uh...you have to leave the service. And I was kind of angry be...not only then, but then later on, as the years unfolded, they changed the rules and regulations. Now, you can work uh, as long as you want. So that's...I remember that very uh, (C - What impact did that have on your life at that point?) Well, I mean I changed uh, (C - Tell me about the changes.) I mean first I had to uh, leave in a hurry, take the airplane and go to XXX and they checked me out. And uh, so then I went back to live in XXX, and later on I picked up things to do and everything was all right.*

If you examine the monologue for the key event, you could argue that the discourse in 'text' is every bit apart of why the individual is angry as the part in 'text', and in fact, the individual could sound angry. However, if you remember the temporal criteria (*first*) you can see how the discourse in 'text' is events leading to the events that are highlighted in 'text', which is what he states made him angry. The discourse in 'text' follows up by briefly expounding on why the events made him angry. Ordering the criteria this way is useful because the subject could be angry during the discourse in 'text', but it would also be appropriate if he was still neutral sounding at that point. However, for the discourse in 'text' and 'text' the subject SHOULD sound angry, and if he does not, that could be meaningful information about the emotionality of this person.

Criterion number 3: Experimenter cues. The last criterion relates to discourse following cues from the experimenter. Although the experimenters have clear guidelines in giving cues, occasionally these guidelines are not strictly followed. Discourse following inappropriate cues by the examiner (see above) is eliminated from discourse analysis. For example, this monologue about a happy event:

Well the last time I really had happiness, my wife had I flew to XXX and we stayed at one of the hotels there and we were able to go so, uh, van and airplane to the various historic, scenic spots in the vicinity. We weren't at all interested in the gambling. It turned us off completely. In addition we have a very happy...very pleasant reunion with two cousins of my wife's and they treated us very graciously. And we also had a reunion with a friend of XXX who lives in XXX. We met her boyfriend and much to our surprise they took us out to dinner and very pleasant time. That's it. (C - Tell us more about that.) XXX and XXX had a show. We were treated to the show and it was really delightful. Uh, they, uh, told us a joke. Let's see if I remember it. Oh, yes uh, it seems like a person and his wife went up to heaven. Oh, I...I forget it. It was a good joke, but I...I louse 'em up. Ok, oh forget that. Um, we saw several shows. We had some fine meals. Uh, we uh...we went to, uh, XXX, but unfortunately, we flew to XXX. we had a nice flight. Lovely flight. 8 passenger twin-engine plane uh, the, uh...unfortunately we weren't dressed for it. It snowed while we were there, despite the fact that it was June. And, uh, we're really cold. We had just enough

time to, uh, glance, down the, uh, canyon and then it was over. We went toXXX.
We went to very...all the very interesting places.

Here the cue simply prompted the subject to give more details in the story and therefore the discourse that follows is included. On the other hand the discourse following the cue in this example about sadness should not be included:

Well I...my mother is uh, uh, 95 years old. And uh, we were living at home, up until last XXX, but uh, it got so it was hard for her to take care of herself. And so we, my sister and myself and her husband, uh, we hadda put her in a home uh, for the aged. She's in a very good uh, mental state, she remembers everything and every time I visit her I...I feel a little sad. Uh, but espec...I felt especially sad uh, last year when we had to bring her to the home. Uh, and this, well, that...that's it. (C - What nursing home is she in? You know, where is it?) Oh, uh, fortunately, uh, the home is about 7, 8, 10 blocks from where I live, so I try to visit her, uh, every uh, every day. It's the XXX Home for the Aged, the XXX division, which is also in XXX, where I live. And my sisters try to...one sister lives close, she tries to see her once or twice...at least twice a week. So that's where I feel...fe...feel sad.

Examples of Cues:

Appropriate, Neutral Cues

Tell me a bit more

Can you say more about that?

Anything else?

Leading, Closed Cues

What happened next?

Where did that happen?

How did this come about?

Appendix C

Rating Procedure Justification: Previous studies (Borod et al., 2000; Canino, 2001; Canino et al., 1999) have measured emotional intensity in a separate session from emotional identification in order to gauge general emotionality independent of specific emotions. This study will use the opposite approach and take these measures at the same time for several reasons. First, raters will not be formally trained (aggregation of rater judgments will be used to generate reliable ratings), and therefore, despite pertinent definitions and an introduction to the rating process, raters will be predominantly relying on their own constructs of what each emotion sounds like and how intensely it is expressed. Raters using their own conceptualization of emotion as the basis of their judgments are important for this study so that relationships between speech acoustics and rater judgments can be made.

In order to establish these relationships concretely, it will be very important to know the frame of reference the rater is using when making their intensity judgment. For example, what is very intense for anger (e.g., fast and loud) is completely the opposite of what is very intense for sad (e.g., slow and soft). So if a rater is going to give an accurate judgment of the emotional intensity of a sample, then they must have a sense of what emotion they are judging. If a rater believes that intensity is only associated with fast and loud, then all the sad samples will be judged to have low intensity, even if for sad it is in fact intensely emotional. Having a rater make their emotional identification judgment first followed by evaluating emotional intensity alleviates this problem.

Furthermore, unlike previous projects that have presented entire monologues in sections (e.g., 15 second increments), this study uses the monologue sampling procedure (previously described) to select the sections of the monologue that are most likely to

reflect the target emotion. This should reduce the emotional variability within each sample and, as such, the ability to gauge general emotionality separate of the target emotion should be greatly reduced – meaning very few samples should have lots of different emotions occurring at the same time, where a general emotionality rating would be useful.

If raters make an intensity rating separately, we will not be sure what their judgment relates to since we are not training them to listen to specific cues to make a judgment. As a result, evaluating the two measures separately loses a lot of experimental control over the intensity judgment and inhibits its utility in the study. In sum, allowing the rater to be conscious about choosing an emotion and then deciding how intense they think it sounds has better construct validity in this case, as we will know what frame of reference the rater is basing their intensity rating on.

The inclusion of neutral and emotional samples together also deserves discussion. Although “Neutral” will NOT be a category option, having samples from the supermarket monologue judged for both emotional category and intensity will help us determine which samples are truly the least emotional later on, as it is unlikely that the neutral monologue will be devoid of any emotional content. Neutral sounding samplings can be differentiated by emotional samples by random category identification by the raters, low or no confidence in that rating, and low emotional intensity.

Random category identification could result for two basic reasons: subjects produced emotional output that reflects some other emotion other than angry, happy, or sad, or the output was essentially neutral and did not fit any emotional category. In both cases, subjects would be forced to guess and as a result both of these situations would result in an equal number of raters picking each emotional category (i.e., angry, happy,

sad). However, "neutral" output should also have low emotional intensity, while emotion that does not fit the three emotional categories should have a higher emotional intensity rating. Therefore, neutral samples from the supermarket monologue should have low emotional identification agreement, low identification confidence, and low emotional intensity. In actuality, some, if not many, of the samples from emotional monologues will be neutral sounding and thus meet these criteria as well. We will also be able to acoustically compare the "neutral" segments from "supermarket" to the "neutral" segments from the "emotional" monologues and see if there are any detectable differences for future reference.

References

- Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *J Affect Disord*, 66(1), 59-69.
- Alpert, M., Rosen, A., Welkowitz, J., Sobin, C., & Borod, J. C. (1989). Vocal acoustic correlates of flat affect in schizophrenia. Similarity to Parkinson's disease and right hemisphere disease and contrast with depression. *Br J Psychiatry Suppl*(4), 51-56.
- Bachorowski, J. A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science*, Vol 8(2), 53-57.
- Bachorowski, J. A., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, Vol 6(4), 219-224.
- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37(13-20).
- Baken, R., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. USA: Singular Thomson Learning.
- Balan, A., & Gandour, J. (1999). Effect of sentence length on the production of linguistic stress by left- and right-hemisphere-damaged patients. *Brain Lang*, 67(2), 73-94.
- Bandler, R., & Shipley, M. T. (1994). Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? *Trends Neurosci* 17(9), 379-389.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol*, 70(3), 614-636.
- Bard, P. (1929). The central representation of the sympathetic system. *Archives of Neurology and Psychiatry*, 22, 230-246.
- Baum, S. R. (1998). The role of fundamental frequency and duration in the perception of linguistic stress by individuals with brain damage. *J Speech Lang Hear Res*, 41(1), 31-40.
- Baum, S. R., & Pell, M. D. (1997). Production of affective and linguistic prosody by brain-damaged patients. *Aphasiology*, Vol 11(2), 177-198.
- Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, Vol 13(8), 581-608.
- Baum, S. R., Pell, M. D., Leonard, C. L., & Gordon, J. K. (1997). The ability of right- and left-hemisphere-damaged individuals to produce and interpret prosodic cues marking phrasal boundaries. *Language and Speech*, Vol 40(4), 313-330.
- Beckman, M. (1997). A typology of spontaneous speech. In Y. Sagisaka & N. Campbell & N. Higuchi (Eds.), *Computing prosody: Computational models for processing spontaneous speech*. New York: Springer.
- Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain*, 114(Pt 3), 1115-1127.
- Blonder, L. X., Pickering, J. E., Heath, R. L., Smith, C. D., & Butler, S. M. (1995). Prosodic characteristics of speech pre- and post-right hemisphere stroke. *Brain Lang*, 51(2), 318-335.
- Bloom, R. L., Borod, J. C., Obler, L. K., & Koff, E. (1990). A preliminary characterization of lexical emotional expression in right and left brain-damaged patients. *Int J Neurosci*, 55(2-4), 71-80.

- Bloom, R. L., Pick, L. H., Borod, J. C., Rorie, K. D., Andelman, F., Obler, L. K., Sliwinski, M., Campbell, A. L., Tweedy, J. R., & Welkowitz, J. (1999). Psychometric aspects of verbal pragmatic ratings. *Brain Lang*, 68(3), 553-565.
- Blumstein, S., & Cooper, W. (1974). Hemisphere processing of intonation contours. *Cortex*, 10, 146-158.
- Boersma, P. (1993). *Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound*. Paper presented at the Proceedings of the Institute of Phonetic Sciences, University of Amsterdam.
- Borod, J., Koff, E., & Buck, R. (1986). The neuropsychology of facial expression in normal and brain-damaged subjects. In N. Fox & R. Davidson (Eds.), *The psychobiology of affective development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Borod, J., Welkowitz, J., & Obler, L. (1992). The New York Emotion Battery. *Unpublished materials, Mount Sinai Medical Center, Department of Neurology, New York*.
- Borod, J., Yecker, S., Moreno, C., Brickman, A., Sliwinski, M., Welkowitz, J., & Alpert, M. (submitted for publication). Changes in posed facial expression of emotion across the adult life span.
- Borod, J., Zgaljardic, D., Tabert, M., & Koff, E. (2001). Asymmetries of emotional communication in normal subjects. In F. Boller & J. G. S. Editor) & G. G. V. Ed.) (Eds.), *Handbook of neuropsychology: Emotional behavior and its disorders*. Oxford, UK: Elsevier Science.
- Borod, J. C. (1992). Interhemispheric and intrahemispheric control of emotion: a focus on unilateral brain damage. *J Consult Clin Psychol*, 60(3), 339-348.
- Borod, J. C. (1993a). Cerebral mechanisms underlying facial, prosodic, and lexical emotional expression: A review of neuropsychological studies and methodological issues. *Neuropsychology*, Vol 7(4), 445-463.
- Borod, J. C. (1993b). Emotion and the brain--anatomy and theory: An introduction to the Special Section. *Neuropsychology*, Vol 7(4), 427-432.
- Borod, J. C., Caron, H. S., & Koff, E. (1981). Asymmetry in positive and negative facial expressions: Sex differences. *Neuropsychologia*, 19(6), 819-824.
- Borod, J. C., & Koff, E. (1990). Lateralization for facial emotional behavior: A methodological perspective. *International Journal of Psychology*, 25(2), 157-177.
- Borod, J. C., Koff, E., Lorch, M. P., & Nicholas, M. (1985). Channels of emotional expression in patients with unilateral brain damage. *Arch Neurol*, 42(4), 345-348.
- Borod, J. C., Koff, E., Yecker, S., Santschi, C., & Schmidt, J. M. (1998). Facial asymmetry during emotional expression: Gender, valence, and measurement technique. *Neuropsychologia*, 36(11), 1209-1215.
- Borod, J. C., & Madigan, N. (2000). Neuropsychology of emotion and emotional disorders: An overview and research directions. In J. Borod (Ed.), *The Neuropsychology of Emotion*. New York: Oxford University Press.
- Borod, J. C., Pick, L. H., Hall, S., Sliwinski, M., Madigan, N., Obler, L. K., Welkowitz, J., Canino, E., Erhan, H. M., Goral, M., Morrison, C., & Tabert, M. (2000). Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition and Emotion*, Vol 14(2), 193-211.
- Boullaud, M. (1825). *Traite clinique et physiologique de l'encephalite ou inflammation du cerveau et de des suites*. Paris: J.B. Bailliere.

- Bowers, D., Bauer, R. M., & Heilman, K. M. (1993). The nonverbal affect lexicon: Theoretical perspectives from neuropsychological studies of affect perception. Special Section: Neuropsychological perspectives on components of emotional processing. *Neuropsychology*, 7(4), 433-444.
- Bowers, D., Coslett, H. B., Bauer, R. M., Speedie, L. J., & Heilman, K. M. (1987). Comprehension of emotional prosody following unilateral hemispheric lesions: processing defect versus distraction defect. *Neuropsychologia*, 25(2), 317-328.
- Bradvik, B., Dravins, C., Holtas, S., Rosen, I., & al., e. (1991). Disturbances of speech prosody following right hemisphere infarcts. *Acta-Neurologica-Scandinavica*, 84(2), 114-126.
- Breitenstein, C., Daum, I., & Ackermann, H. (1998). Emotional processing following cortical and subcortical brain damage: Contribution of the fronto-striatal circuitry. *Behavioural Neurology*, Vol 11(1), 29-42.
- Breitenstein, C., Van Lancker, D., Daum, I., & Waters, C. H. (2001). Impaired perception of vocal emotions in Parkinson's disease: influence of speech time processing and executive functioning. *Brain Cogn*, 45(2), 277-314.
- Broca, P. (1861). Remarques sur le Siege de la faculte du langage articule, suivies d'une observation d'aphemie (perte de la parole). *Bulletins de la Societe Anatomicque*, 36, 330-357.
- Bryan, K. (1989). Language prosody and the right hemisphere. *Aphasiology*, 3, 285-299.
- Bryden, M. P. (1982). *Laterality*. San Diego, CA: Academic Press.
- Bryden, M. P., & Ley, R. G. (1983). Right-hemisphere involvement in the perception and expression of emotion in normal humans. In K. M. Heilman & P. Satz (Eds.), *Neuropsychology of human emotion* (pp. 6-44). New York: Guilford Press.
- Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jancke, L. (2000). Recognition of emotional prosody and verbal components of spoken language: an fMRI study. *Brain Res Cogn Brain Res*, 9(3), 227-238.
- Buck, R. (1984). *The communication of emotion*. New York: Guilford Press.
- Buck, R., & Duffy, R. J. (1980). Nonverbal communication of affect in brain-damaged subjects. *Cortex*, 16(351-362).
- Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annu Rev Psychol*(Null), 50191-50214.
- Cancelliere, A. E., & Kertesz, A. (1990). Lesion localization in acquired deficits of emotional expression and comprehension. *Brain-and-Cognition*, 13(2), 133-147.
- Canino, E. (2001). *Unpublished dissertation*. Queens College, Flushing, NY.
- Canino, E., Borod, J. C., Madigan, N., Tabert, M. H., & Schmidt, J. M. (1999). Development of procedures for rating posed emotional expressions across facial, prosodic, and lexical channels. *Percept Mot Skills*, 89(1), 57-71.
- Carlson, N. R. (2001). *Physiology of Behavior* (7th ed.). Boston: Allyn and Bacon.
- Coren, S., Porac, C., & Duncan, P. (1979). A behaviorally validated self-report inventory to assess four types of lateral preferences. *Journal of Clinical Neuropsychology*, 1, 55-64.
- Cummings, J. L. (1995). Anatomic and behavioral aspects of frontal-subcortical circuits. *Ann N Y Acad Sci*, 15(Null), 7691-7613.
- Davidson, R. (1984). Affect, cognition, and hemispheric specialization. In C. E. Izard & J. Kagan & R. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 320-365). Cambridge, England: Cambridge University Press.

- Davidson, R. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science*, 3, 39-43.
- Davidson, R. J. (1993). Parsing affective space: Perspectives from neuropsychology and psychophysiology. *Neuropsychology*, Vol 7(4), 464-475.
- Davidson, R. J. (1998). Cerebral Asymmetry, Emotion, and Affective Style. In R. J. Davidson & K. Hugdahl (Eds.), *Brain Asymmetry* (pp. 361-387). Cambridge, MA: The MIT Press.
- Davidson, R. J., Abercrombie, H., Nitschke, J. B., & Putnam, K. (1999). Regional brain function, emotion and disorders of emotion. *Curr Opin Neurobiol*, 9(2), 228-234.
- Dykstra, K., Gandour, J., & Stark, R. E. (1995). Disruption of prosody after frontal lobe seizures in the non-dominant hemisphere. *Aphasiology*, 9(5), 453-476.
- Edmondson, J. A., Chan, J. I., Seibert, G. B., & Ross, E. D. (1987). The effect of right-brain damage on acoustical measures of affective prosody in Taiwanese patients. *Journal-of-Phonetics*, 15(3), 219-233.
- Egelko, S., Gordon, W., Hibbard, M., Diller, L., Lieberman, A., Holliday, R., Ragnarsson, K., Shaver, M., & Orazem, J. (1988). Relationship among CT scans, neurological exam, and neuropsychological test performance in right-hemisphere-damaged stroke patients. *Journal of Clinical and Experimental Neuropsychology*, 10, 539-564.
- Ekman, P. (1973). Universal facial expressions in emotion. *Studia Psychologica*, 15(2), 140-147.
- Ekman, P. (1992). Are there basic emotions? *Psychol Rev*, 99(3), 550-553.
- Ekman, P., & Davidson, R. J. (Eds.). (1994). *The nature of emotion: Fundamental questions*.
- Ekman, P., & Friesen, W. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.
- Ekman, P., Friesen, W., O'Sullivan, M., & Scherer, K. (1980). Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology*, 38, 207-277.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*.
- Ellis, L. (2001). The biosocial female choice theory of social stratification. *Soc Biol*, 48(3-4), 298-320.
- Endicott, J., & Spitzer, R. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry*, 35, 837-844.
- Epstein, S. (1980). The stability of behavior: Implications for psychological research. *American Psychologist*, 35(9), 790-806.
- Erhan, H., Borod, J. C., Tenke, C. E., & Bruder, G. E. (1998). Identification of emotion in a dichotic listening task: event-related brain potential and behavioral findings. *Brain Cogn*, 37(2), 286-307.
- Esposito, A., Demeurisse, G., Alberti, B., & Fabbro, F. (1999). Complete mutism after midbrain periaqueductal gray lesion. *Neuroreport*, 10(4), 681-685.
- Farkas, E., Jansen, A. S., & Loewy, A. D. (1997). Periaqueductal gray matter projection to vagal preganglionic neurons and the nucleus tractus solitarius. *Brain Res*, 764(1-2), 257-261.
- Fisk, G. D., & Wyss, J. M. (2000). Descending projections of infralimbic cortex that mediate stimulation-evoked changes in arterial pressure. *Brain Res*, 859(1), 83-95.

- Fitzsimons, M., Sheahan, N., & Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: Implications for clinical studies. *Brain and Language*, *Vol 78*(1), 94-108.
- Gainotti, G. (2001). Neuropsychological Theories of Emotion. *The Neuropsychology of Emotion*, Null(Null), Null.
- Gainotti, G., Caltagirone, C., & Zoccolotti, P. (1993). Left/right and cortical/subcortical dichotomies in the neuropsychological study of human emotions. *Cognition and Emotion*, *7*, 71-93.
- Gandour, J., Larsen, J., Dechongkit, S., Ponglorpisit, S., & Khunadorn, F. (1995). Speech prosody in affective contexts in Thai patients with right hemisphere lesions. *Brain Lang*, *51*(3), 422-443.
- Garcia Toro, M., Talavera, J. A., Saiz Ruiz, J., & Gonzalez, A. (2000). Prosody impairment in depression measured through acoustic analysis. *J Nerv Ment Dis*, *188*(12), 824-829.
- George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., Herscovitch, P., & Post, R. M. (1996). Understanding emotional prosody activates right hemisphere regions. *Arch Neurol*, *53*(7), 665-670.
- Goldberg, E., & Costa, L. D. (1981). Hemisphere differences in the acquisition and use of descriptive systems. *Brain Lang*, *14*(1), 144-173.
- Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders*. Philadelphia, PA: Lea & Febiger.
- Gorelick, P. B., & Ross, E. D. (1987). The aprosodias: further functional-anatomical evidence for the organisation of affective language in the right hemisphere. *J Neurol Neurosurg Psychiatry*, *50*(5), 553-560.
- Gur, R. C., Packer, I. K., Hungerbuhler, J. P., Reivich, M., Obrist, W. D., Amarnek, W. S., & Sackeim, H. A. (1980). Differences in the distribution of gray and white matter in human cerebral hemispheres. *Science*, *207*(4436), 1226-1228.
- Harrington, A. (1998). Unfinished business: Models of laterality in the nineteenth century. In R. Davidson & K. Hugdahl (Eds.), *Brain Asymmetry* (pp. 3-27). Massachusetts: MIT Press.
- Heilman, K. M., Blonder, L. X., Bowers, D., & Crucian, G. P. (2000). Neurological disorders and emotional dysfunction. In J. C. Borod (Ed.), *The neuropsychology of emotion*. New York: Oxford University Press.
- Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology*, *34*(7), 917-921.
- Heilman, K. M., & Gilmore, R. L. (1998). Cortical influences in emotion. *J Clin Neurophysiol*, *15*(5), 409-423.
- Heilman, K. M., Scholes, R., & Watson, R. (1975). Auditory affective agnosia: Disturbed comprehension of affective speech. *Journal of Neurology, Neurosurgery and Psychiatry*, *38*, 69-72.
- Heller, W. (1990). The neuropsychology of emotion: Developmental patterns and implications for psychopathology. In N. Stein & B. L. Leventhal & T. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 167-211). Hillsdale, NJ: Lawrence Erlbaum.
- Hird, K., & Kirsner, K. (1993). Dysprosody following acquired neurogenic impairment. *Brain Lang*, *45*(1), 46-60.
- Hollingshead, A. (1977). Four-factor index of social status. *Unpublished manuscript*. Yale University Department of Sociology, New Haven, CT.

- Holstege, G., & Ehling, T. (1996). Two Motor Systems Involved in the Production of Speech. *Vocal Fold Physiology: Controlling Complexity and Chaos*, Null(Null), Null.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: John Wiley & Sons, Inc.
- Hughes, C. P., Chan, J. L., & Su, M. S. (1983). Aprosodia in Chinese patients with right cerebral hemisphere lesions. *Arch Neurol*, 40(12), 732-736.
- Izard, C. (1977). *Human emotions*. New York: Plenum.
- Jackson, J. H. (1874). On the nature of the duality of the brain. *The medical Press & Circular*, 1, 41-44.
- Jackson, J. H. (1880). On affections of speech from disease of the brain. *Brain*, 2, 203-222.
- Jackson, J. H. (1915). On affections of speech from diseases of the brain. *Brain*, 38, 101-186.
- Jurgens, U. (1994). The role of the periaqueductal grey in vocal behaviour. *Behav Brain Res*, 62(2), 107-117.
- Jurgens, U., & Zwirner, P. (1996). The role of the periaqueductal grey in limbic and neocortical vocal fold control. *Neuroreport*, 7(18), 2921-2923.
- Kaplan, P. S., Bachorowski, J. A., Smoski, M. J., & Hudenko, W. J. (2002). Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychol Sci*, 13(3), 268-271.
- Kaufer, D., & Lewis, D. (1996). Frontal lobe anatomy and cortical connectivity. In B. Miller & J. Cummings (Eds.), *The human frontal lobes* (pp. 27-44). New York: The Guilford Press.
- Lane, R., Reiman, E., Ahern, G., Schwartz, G., & Davidson, R. (1997). Neuroanatomical correlates of happiness, sadness, and disgust. *Am J Psychiatry*, 154(7), 926-933.
- Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A., & Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness. Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *J Cogn Neurosci*, 10(4), 525-535.
- Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., & Schwartz, G. E. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, 35(11), 1437-1444.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261-273.
- Langer, S. L., Pettigrew, L. C., Wilson, J. F., & Blonder, L. X. (1998). Personality and social competency following unilateral stroke. *J Int Neuropsychol Soc*, 4(5), 447-455.
- Lebrun, Y., Lessinnes, A., De Vresse, L., & Leleux, C. (1985). Dysprosody and the non-dominant hemisphere. *Language Science*, 7(41-52).
- Leentjens, A. F., Wilaert, S. M., van Harskamp, F., & Wilmkink, F. W. (1998). Disturbances of affective prosody in patients with schizophrenia: a cross sectional study. *J Neurol Neurosurg Psychiatry*, 64(3), 375-378.
- Lewis, J. A. (2002). *Social Influences on Female Speaker's Pitch*. Unpublished Dissertation, University of California, Berkeley.
- Ley, R. G., & Bryden, M. P. (1982). A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content. *Brain Cogn*, 1(1), 3-9.

- Martin, C. C., Borod, J. C., Alpert, M., Brozgold, A., & Welkowitz, J. (1990). Spontaneous expression of facial emotion in schizophrenic and right-brain-damaged patients. *J Commun Disord*, 23(4-5), 287-301.
- Mattis, S. (1988). *Dementia Rating Scale*. Odessa, FL: Psychological Assessment Resources, Inc.
- Mega, M. S., & Cummings, J. L. (1994). Frontal-subcortical circuits and neuropsychiatric disorders. *J Neuropsychiatry Clin Neurosci*, 6(4), 358-370.
- Morris, J. S., Scott, S. K., & Dolan, R. J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia*, 37(10), 1155-1163.
- Ostrove, J. M., Simpson, T., & Gardner, H. (1990). Beyond scripts: a note on the capacity of right hemisphere-damaged patients to process social and emotional content. *Brain Cogn*, 12(1), 144-154.
- Ouellette, G. P., & Baum, S. R. (1994). Acoustic analysis of prosodic cues in left- and right-hemisphere-damaged patients. *Aphasiology*, 8(3), 257-283.
- Panksepp, J. (1992). A critical role for "affective neuroscience" in resolving what is basic about basic emotions. *Psychol Rev*, 99(3), 554-560.
- Panksepp, J. (1998). *Affective Neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Papez, J. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 79, 217-224.
- Pell, M. D. (1996). On the receptive prosodic loss in Parkinson's disease. *Cortex*, Vol 32(4), 693-704.
- Pell, M. D. (1998). Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. *Neuropsychologia*, Vol 36(8), 701-715.
- Pell, M. D. (1999a). Fundamental frequency encoding of linguistic and emotional prosody by right hemisphere-damaged speakers. *Brain and Language*, Vol 69(2), 161-192.
- Pell, M. D. (1999b). Some acoustic correlates of perceptually "flat affect" in right-hemisphere-damaged speakers. *Brain and Cognition*, Vol 40(1), 219-223.
- Pell, M. D. (1999c). The temporal organization of affective and non-affective speech in patients with right-hemisphere infarcts. *Cortex*, Vol 35(4), 455-477.
- Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *J Acoust Soc Am*, 109(4), 1668-1680.
- Pell, M. D. (2002). Evaluation of nonverbal emotion in face and voice: some preliminary findings on a new battery of tests. *Brain Cogn*, 48(2-3), 499-504.
- Pell, M. D., & Baum, S. R. (1997). Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain Lang*, 57(2), 195-214.
- Pihan, H. (2001). *Emotional valence and intonational sentence mode conveyed by direction of pitch contour*. Unpublished manuscript, New York.
- Pihan, H., Altenmuller, E., Hertrich, I., & Ackermann, H. (2000). Cortical activation patterns of affective speech processing depend on concurrent demands on the subvocal rehearsal system. A DC-potential study. *Brain*, 123(Pt 11), 2338-2349.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus*. New York: Springer.
- Pittam, J., & Scherer, K. (1993). Vocal expression and communication of emotion. In M. Lewis & J. Haviland (Eds.). *The handbook of emotions* (pp. 185-197). New York: Guilford.

- Plutchik, R. (1962). *The emotions: facts, theories, and a new model*.
- Plutchik, R., & Kellerman, H. (1980). *Emotion: Theory, research, and experience (Vol. 1)*. New York: Academic Press.
- Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measure of jitter. *J Speech Hear Res*, 38(1), 26-32.
- Robinson, R. G., & Manes, F. (2000). Elation, mania, and mood disorders: Evidence from neurological disease. In B. J. (Ed.), *The Neuropsychology of Emotion*. New York: Oxford University Press.
- Rolls, E. T. (2000). Precis of The brain and emotion. *Behav Brain Sci*, 23(2), 177-191.
- Ross, E. D. (1993). Nonverbal aspects of language. *Neurol Clin*, 11(1), 9-23.
- Ross, E. D., Edmondson, J. A., & Seibert, G. B. (1986). The effect of affect on various acoustic measures of prosody in tone and non-tone languages: A comparison based on computer analysis of voice. *Journal of Phonetics*, Vol 14(2), 283-302.
- Ross, E. D., Edmondson, J. A., Seibert, G. B., & Homan, R. W. (1988). Acoustic analysis of affective prosody during right-sided Wada Test: a within-subjects verification of the right hemisphere's role in language. *Brain Lang*, 33(1), 128-145.
- Ross, E. D., Homan, R. W., & Buck, R. (1994). Differential hemispheric lateralization of primary and social emotions: Implications for developing a comprehensive neurology for emotions, repression, and the subconscious. *Null*, Vol 7(1), 1-19.
- Ross, E. D., & Mesulam, M. M. (1979). Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Archives-of-Neurology*, 36(3), 144-148.
- Ross, E. D., Thompson, R. D., & Yenkosky, J. (1997). Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang*, 56(1), 27-54.
- Ryalls, J., Joannette, Y., & Feldman, L. (1987). An acoustic comparison of normal and right-hemisphere-damaged speech prosody. *Cortex*, 23(4), 685-694.
- Sackeim, H. A., Greenberg, M. S., Weiman, A. L., Gur, R. C., Hungerbuhler, J. P., & Geschwind, N. (1982). Hemispheric asymmetry in the expression of positive and negative emotions. Neurologic evidence. *Arch Neurol*, 39(4), 210-218.
- Scherer, K. (2001). Psychological Models of Emotion. In J. Borod (Ed.), *The Neuropsychology of Emotion* (pp. 137-162). New York.
- Scherer, K. R. (2000). Psychological Models of Emotion. In J. C. Borod (Ed.), *The Neuropsychology of Emotion* (Vol. Null, pp. Null): 1651.
- Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion*, Vol 15(2), 123-148.
- Schirmer, A., Alter, K., Kotz, S. A., & Friederici, A. D. (2001). Lateralization of prosody during language production: a lesion study. *Brain Lang*, 76(1), 1-17.
- Semenza, C., Pasini, M., Zettin, M., Tonin, P., & Portolan, P. (1986). Right hemisphere patients' judgments on emotions. *Acta Neurologica Scandinavica*, 74, 43-50.
- Semmes, J. (1968). Hemispheric Specialization: A Possible Clue to Mechanism. *Neuropsychologia*, 6(1), 11-26.
- Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain Lang*, 25(1), 19-36.
- Silberman, E. K., & Weingartner, H. (1986). Hemispheric lateralization of functions related to emotion. *Brain Cogn*, 5(3), 322-353.

- Sobin, C., & Alpert, M. (1999). Emotion in speech: the acoustic attributes of fear, anger, sadness, and joy. *J Psycholinguist Res*, 28(4), 347-365.
- Speilman, J. (2002). Acoustics. In M. Schmidt (Ed.). New York.
- Springer, S., & Deutsch, G. (1998). *Left brain right brain: Perspectives from cognitive neuroscience* (5th ed.). New York: W. H. Freeman and Company.
- Starkstein, S. E., Federoff, J. P., Price, T. R., Leiguarda, R. C., & Robinson, R. G. (1994). Neuropsychological and neuroradiologic correlates of emotional prosody comprehension. *Neurology*, 44(3 Pt 1), 515-522.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics* (3rd ed.). New York: HarperCollins.
- Thatcher, R. W., Krause, P. J., & Hrybyk, M. (1986). Cortico-cortical associations and EEG coherence: a two-compartmental model. *Electroencephalogr Clin Neurophysiol*, 64(2), 123-143.
- Tompkins, C. A., & Flowers, C. R. (1985). Perception of emotional intonation by brain-damaged adults: the influence of task processing levels. *J Speech Hear Res*, 28(4), 527-538.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon Bull Rev*, 9(2), 335-340.
- Trauner, D. A., Ballantyne, A., Friedland, S., & Chase, C. (1996). Disorders of affective and linguistic prosody in children after early unilateral brain damage. *Ann Neurol*, 39(3), 361-367.
- Trotman, S. C., & Hammond, G. R. (1989). Lateral asymmetry of the scalp distribution of somatosensory evoked potential amplitude. *Brain Cogn*, 10(1), 132-147.
- Tucker, D., Watson, R., & Heilman, K. (1977). Discrimination and evocation of affective intoned speech in patients with right parietal disease. *Neurology*, 27, 947-950.
- Tucker, D. M. (1981). Lateral brain function, emotion, and conceptualization. *Psychol Bull*, 89(1), 19-46.
- Tucker, D. M. (1989). Neural and psychological maturation in a social context. *Cicchetti, Dante (Ed). (1989). The emergence of a discipline: Rochester Symposium on Developmental Psychopathology, Vol. 1. (pp. 69 88). . xii. 310 pp.SEE BOOK.*
- Tucker, D. M., Roth, D. L., & Bair, T. B. (1986). Functional connections among cortical regions: topography of EEG coherence. *Electroencephalogr Clin Neurophysiol*, 63(3), 242-250.
- Twist, D. J., Squires, N. K., Spielholz, N. I., & Silverglide, R. (1991). Event-related potentials in disorders of prosodic and semantic linguistic processing: I. *Null*, 4(4), 281-304.
- Van Bockstaele, E. J., Bajic, D., Proudfit, H., & Valentino, R. J. (2001). Topographic architecture of stress-related pathways targeting the noradrenergic locus coeruleus. *Physiol Behav*, 73(3), 273-283.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *International Journal of Human Communication*, 13(227-277).
- Van Lancker, D., & Sidtis, J. J. (1992). The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. *Journal-of-Speech-and-Hearing-Research*, 35(5), 963-970.
- Wechsler, D. (1981). *WAIS-R manual: Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.
- Weintraub, S., Mesulam, M. M., & Kramer, L. (1981). Disturbances in prosody. A right-hemisphere contribution to language. *Arch Neurol*, 38(12), 742-744.

- Welkowitz, J., Bond, R., & Zelano, J. (1990). An automated system for the analyses of temporal speech patterns: Description of the hardware and software. *Journal of Communication Disorders, 23*, 347-364.
- Woodward, S. (1988). An anatomical model of hemisphere asymmetry [Abstract]. *Journal of Clinical and Experimental Neuropsychology, 10*(68).
- Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*(846-849).
- Zgaljardic, D. (2000). Recovery of emotional perception in brain-damaged patients. *Unpublished master's thesis. Queens College, Psychology Department, New York.*
- Zurif, E., & Mendelsohn, M. (1972). Hemisphere specialization for the perception of speech sounds: The influence of intonation and structure. *Perception & Psychophysics, 11*, 329-332.