Fundamental Frequency Encoding of Linguistic and Emotional Prosody by Right Hemisphere-Damaged Speakers

Marc D. Pell

School of Communication Sciences and Disorders, McGill University, Montréal, Canada

To illuminate the nature of the right hemisphere’s involvement in expressive prosodic functions, a story completion task was administered to matched groups of right hemisphere-damaged (RHD) and nonneurological control subjects. Utterances which simultaneously specified three prosodic distinctions (emphatic stress, sentence modality, emotional tone) were elicited from each subject group and then subjected to acoustic analysis to examine various fundamental frequency (F₀) attributes of the stimuli. Results indicated that RHD speakers tended to produce F₀ patterns that resembled normal productions in overall shape, but with significantly less F₀ variation. The RHD patients were also less reliable than normal speakers at transmitting emphasis or emotional contrasts when judged from the listener’s perspective. Examination of the results across a wide variety of stimulus types pointed to a deficit in successfully implementing continuous aspects of F₀ patterns following right hemisphere insult. © 1999 Academic Press

INTRODUCTION

A disturbed capacity to modulate prosodic attributes of speech (principally, fundamental frequency, timing, and intensity) has been linked to focal cortical lesions of both the left and right cerebral hemispheres (Baum & Pell, 1997; Baum, Pell, Leonard, & Gordon, 1997; Behrens, 1988, 1989; Blonder, Pickering, Heath, Smith, & Butler, 1995; Cancelliere & Kertesz, 1990; Cooper, Soares, Nicol, Michelow, & Goloskie, 1984; Danly, Cooper, & Shapiro, 1984; I am thankful to Drs. Shari Baum, Rachel Mayberry, and John Ryalls for comments on an earlier version of the manuscript and to Marie Leconte, Anita Shuper, and Denis Côté for help in data collection and manuscript preparation. The commitment of the brain-damaged individuals to this research is particularly noted. This work was supported by funds granted to the author through the Fonds de la Recherche en Santé du Québec (FRSQ) and the McGill University Research Development Fund. Address correspondence and reprint requests to Marc D. Pell, McGill University, School of Communication Sciences & Disorders, 1266 Pine Ave West, Montréal, Québec, H3G 1A8 Canada. E-mail: czbm@musica.mcgill.ca. 0093-934X/99 $30.00 Copyright © 1999 by Academic Press All rights of reproduction in any form reserved.
Defects in regulating prosodic attributes of speech impose severe limitations on the communicative proficiency of many brain-damaged individuals. Such impairments may hamper the ability to mark aspects of linguistic structure, such as lexical or emphatic stress, or to use prosody to distinguish the declarative from the interrogative speaking mode, for example (Behrens, 1988, 1989; Cooper et al., 1984; Shapiro & Danly, 1985; Weintraub et al., 1981). Difficulties implementing the same intonational parameters may simultaneously obscure vital cues about the patient’s state of arousal or *emotional* disposition, leading to confusion on the part of the listener as to whether the speaker is happy, sad, or angry (Baum & Pell, 1997; Edmondson, Chan, Seibert, & Ross, 1987; Gandour et al., 1995; Tucker, Watson, & Heilman, 1977).

The dual (i.e., linguistic or emotional) behavioral role specified by prosodic features, although not uniformly accepted (e.g., Bolinger, 1986; Crystal, 1969; Ladd, 1980), has proven a fertile concept for researchers in explaining brain–prosody relationships, stimulating a number of hypotheses (Emmorey, 1987; Gandour et al., 1995, 1992; Ross, 1981; Van Lancker, 1980). The dominant view emerging from this research ascribes a privileged role to the right hemisphere in modulating emotional attributes of prosodic patterns and a dominant role to the left hemisphere in encoding linguistic parameters of the same contours (Van Lancker, 1980). This “functional” hypothesis, which received much initial support from clinical impressions of restricted emotional inflection or “flattened affect” in right hemisphere-damaged (RHD) speakers (Ross, 1981; Ross & Mesulam, 1979), has been more rigorously advanced by acoustic studies of RHD and left hemisphere-damaged (LHD) speakers producing both linguistic and emotional stimuli (Gandour et al., 1995, 1992). Conflicting patterns of functional lateralization have been noted, however, including right hemisphere jurisdiction over both linguistic and emotional attributes of prosodic contours (Shapiro & Danly, 1985) or bilateral control for just linguistic–prosodic stimuli (Heilman, Bowers, Speedic, & Coslett, 1984; Pell & Baum, 1997a). Discrepancies among these studies are not easily reconciled owing to methodological irregularities, including the types of prosodic structures investigated (emotional or linguistic prosody in isolation or a combination of the two), the manner in which language samples were obtained (repetition or reading paradigm vs. spontaneous speech), and/or the clinical characteristics of the patient samples assessed.

A somewhat different conception of how the brain may formulate prosodic messages focuses on potential hemispheric asymmetries in the ability to en-
code individual acoustic parameters of intonation contours. More precisely, the right hemisphere may be more suited to process fundamental frequency ($F_0$) attributes of prosodic contours and the left hemisphere may be specialized to regulate temporal properties of the same forms (Cooper, 1981; Robin, Tranel, & Damasio, 1990; Van Lancker & Sidtis, 1992). Such a dichotomy obtains support from studies of spectral and temporal processing of auditory patterns during nonlinguistic tasks (e.g., Carmon & Nachshon, 1971; Sidtis & Volpe, 1988). Although several studies have emphasized the importance of individual acoustic parameters to prosody lateralization (Baum & Pell, 1997; Dykstra, Gandour, & Stark, 1995; Emmorey, 1987; Hird & Kirchner, 1993; Ouellette & Baum, 1993), the patterns detected do not always correspond precisely to the dichotomy outlined above or may be contingent on other characteristics of the stimuli, such as the size of the processing domain (Behrens, 1989; Emmorey, 1987).

It is noteworthy that a dichotomy based on acoustic–perceptual features—if shown to be a true reflection of hemispheric involvement in prosodic functions—cuts across hypothetical divisions in the behavior of prosodic stimuli. Indeed, there is ample evidence that $F_0$ and duration constitute the primary cues to both linguistic– and emotional–prosodic messages in speech production (Bolinger, 1955; Klatt, 1976; Scherer, 1986). However, diverging predictions with respect to how functional and acoustic parameters engage the two hemispheres do not preclude the possibility that hemispheric performance is sensitive to both types of attributes in an independent, and perhaps at times interactive, manner. This prospect has not been addressed in previous expressive studies of prosody in brain-damaged speakers (but see Pell, 1998, for recent data on receptive abilities). If demonstrated, such an interaction might bolster recent claims that prosodic functions are multifaceted in nature and cannot be qualified by a strict right hemisphere substrate (Van Lancker & Sidtis, 1992).

Certainly, neurolinguistic investigations of prosody are conspicuous in emphasizing a right hemisphere substrate for at least some aspects of prosodic encoding. This pattern of hemispheric performance does not obtain from studies of linguistic (segmental) processing, for which right hemisphere mechanisms are known to be minimally implicated. Detailed analysis of RHD patients, therefore, holds the potential of illuminating unique right hemisphere subsystems dedicated to prosody (or aspects thereof) which are not directly involved in generating segmental aspects of a verbal message. Investigating the manner in which focal right-brain injury affects the transmission of multiple prosodic cues—both linguistic and affective in nature and encompassing different processing domains—may begin to illuminate the precise nature of the intact right hemisphere’s capacity for prosody production. Indeed, understanding the interactive influence of various prosodic distinctions (certainly a natural characteristic of speech production) on RHD speakers constitutes one of the primary objectives of the current investigation.
of fundamental frequency encoding and that of a companion report on tem-portal encoding by the same speakers (Pell, in press (b)).

In spoken English, the primacy of fundamental frequency in expressing emotional content and in transmitting many linguistic distinctions (such as word stress) is seldom disputed (Cosmides, 1983; Lieberman & Michaels, 1962; Pell, 1998; Scherer, 1986). These forms are typically marked through modifications in the height and/or variation of F0 over relevant segmental units (Frick, 1985; Ladd, Silverman, Tolkmitt, Bergmann, & Scherer, 1985; Scherer, 1986). Alternatively, the modulation of F0 to encode prosodic meaning may be viewed as operating categorically (e.g., detection of a pitch accent on a word signals the presence of linguistic focus) or as operating within a gradient of acoustic change (e.g., “happy” tends to be expressed with greater F0 variation than “sad” by most speakers). In the latter case, meaning is not assigned in an all-or-none fashion but in more relative terms (Huttar, 1968; Ladd et al., 1985; Scherer, 1986). It is noteworthy that the continuous properties of F0 contours, although strongly associated with the speaker’s overall level of activation or arousal (i.e., emotional prosody), may also reflect differences in the extent to which focused words are marked for emphasis or the extent to which interrogation is being conveyed, etc. (e.g., Bolinger, 1961).

Recent reports in the literature have begun to speculate about a relationship between right hemisphere dysfunction and an incapacity to regulate only graded, or continuous, aspects of intonation contours (Blonder et al., 1995; Pell, 1997b). Such a link cannot easily be discerned from previous studies of prosody, as researchers have preferred to interpret their results in light of the assumed behavioral properties of the stimuli, rather than the manner in which those forms are encoded and decoded in the speech signal. Accordingly, the present inquiry attempts to additionally monitor the production of both continuous and categorical attributes of F0 contours in the expression of linguistic focus, sentence modality contrasts (declarative/interrogative), and emotion by RHD and normal speakers. These results may shed new light on how the acoustic representation of prosodic meaning within the speech signal influences the right hemisphere’s commitment to prosodic operations in speech production.

METHOD

Subjects

Experimental subjects included 10 individuals with unilateral right hemisphere damage and 10 healthy control subjects without neurological dysfunction. All subjects were right-handed native speakers of Canadian English (R8 reported favoring the left hand at birth) with no premorbid history of neurologic episodes, psychiatric illness, or substance abuse. All subjects displayed acceptable hearing levels at 0.5, 1, and 2 kHz following a pure-tone air-conduction screening of both ears (inclusion criteria: 30 dB HL at each frequency, for the better ear).
The RHD and NC groups were balanced for mean age in years (RHD = 64.3; NC = 66.1) and were closely balanced for gender.

Brain-damaged participants in the study were approached at one of three Montreal-area rehabilitation centers following a single embolic or thrombotic event in the distribution of the right middle cerebral artery. All lesions were documented by CT scan; this information was derived from radiologic reports in patients’ medical records. Evidence of bilateral or brainstem lesions, or the presence of a condition associated with cognitive deterioration (e.g., Parkinson’s disease), served as grounds for exclusion from the study. None of the RHD subjects presented signs of aphasic or dysarthric deficit as determined by an aphasia screening and through clinical observation. Hemispatial visual neglect was identified in four RHD patients following administration of the Bells Test (Gauthier, Dehaut, & Joanette, 1989). Other behavioral measures included an evaluation of discourse inferencing abilities, figurative language comprehension, and emotional prosody discrimination and recognition; deficits on these subtests are reported in Table 1, where observed. Testing was conducted at least 3 months postonset of stroke (X = 36 months, range = 7–83) following patient discharge from the rehabilitation facility.

Materials

Following previous work on normal speech attributes (Cooper, Eady, & Mueller, 1985a; Eady & Cooper, 1986; Eady, Cooper, Klouda, Mueller, & Lotts, 1986; McRoberts, Studdert-Kennedy, & Shankweiler, 1995), two short sentences, each adhering to the basic subject–verb–object syntactic ordering of English, served as test items:

(1) Barry took the sailboat
(2) Mary sold the teapot

Each sentence was composed of high frequency content words matched across the two stimuli for syllable length and stress placement. When spoken, each item was constructed to permit multiple variations in the prosodic form of the utterance as an index of three factors: location of sentential focus (initial, final, none); linguistic ‘‘modality’’ (statement, question); and emotional tone (angry, sad, happy, neutral). Given the potential interactive influence of these prosodic variables on speech programming, an exhaustive combination of the three factors was elicited from each subject for each test item. This resulted in a corpus of 24 productions of each item (3 focus locations × 2 modalities × 4 emotions) or 48 productions per speaker in total. It should be highlighted that although 24 unique productions of each test item were elicited when considered at the suprasegmental level, all productions contained identical segmental content, permitting controlled study of the acoustic parameters employed by each group in the communication of prosodic meaning.

To elicit specific constellations of prosodic features differing in focus location, sentence modality, and emotion, a story completion paradigm was adopted (Behrens, 1989; Cooper et al., 1985a). Short scenarios of three to four sentences in length were constructed to precede each trial, biasing the subjects’ reading of the stimulus. For example, the scenario preceding a neutral, declarative reading of Mary sold the teapot without sentence focus (i.e., [no focus, declarative, neutral]) was the following:

You are holding a garage sale at your house with the help of some friends. After the sale, someone tells you that Mary sold the teapot. When another friend asks you what happened, you say:

[MARY SOLD THE TEAPOT]

The content of the scenarios was modified or elaborated to facilitate the various prosodic “targets” as follows: to elicit focus distinctions, the information to be used contrastively (e.g., Mary, teapot) was not presented as “given” information in the prime. Differences in linguistic modality were evoked by requiring the speaker to address a fictional character (i.e., the
### TABLE 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Age</th>
<th>Postonset</th>
<th>Lesion site</th>
<th>Major clinical signs</th>
<th>VN</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>F</td>
<td>54</td>
<td>66</td>
<td>(R) Posterior</td>
<td>(L) Hemiparesis, “flat affect”</td>
<td>−</td>
</tr>
<tr>
<td>R2</td>
<td>M</td>
<td>59</td>
<td>31</td>
<td>(R) Temporoparietal</td>
<td>(L) Hemiparesis, “inappropriate mood,” impaired inferencing, comprehension of figurative language</td>
<td>+</td>
</tr>
<tr>
<td>R3</td>
<td>M</td>
<td>74</td>
<td>48</td>
<td>(R) Parietal</td>
<td>(L) Hemiparesis, impaired comprehension of figurative language</td>
<td>−</td>
</tr>
<tr>
<td>R4</td>
<td>F</td>
<td>61</td>
<td>8</td>
<td>(R) Basal ganglia</td>
<td>(L) Hemiparesis, “flat affect”</td>
<td>−</td>
</tr>
<tr>
<td>R5</td>
<td>F</td>
<td>29</td>
<td>7</td>
<td>(R) MCA</td>
<td>(L) Hemiplegia</td>
<td>−</td>
</tr>
<tr>
<td>R6</td>
<td>F</td>
<td>82</td>
<td>40</td>
<td>(R) Temporal</td>
<td>(L) Hemiplegia, impaired recognition of emotional prosody</td>
<td>+</td>
</tr>
<tr>
<td>R7</td>
<td>M</td>
<td>69</td>
<td>59</td>
<td>(R) Temporoparietal</td>
<td>(L) Hemiparesis, impaired recognition of emotional prosody</td>
<td>−</td>
</tr>
<tr>
<td>R8</td>
<td>F</td>
<td>87</td>
<td>83</td>
<td>(R) MCA</td>
<td>(L) Hemiparesis, impaired inferencing/comprehension of figurative language</td>
<td>−</td>
</tr>
<tr>
<td>R9</td>
<td>F</td>
<td>62</td>
<td>10</td>
<td>(R) External capsule</td>
<td>(L) Hemiparesis, “flat affect”, impaired inferencing/comprehension of figurative language/ emotional prosody recognition</td>
<td>+</td>
</tr>
<tr>
<td>R10</td>
<td>F</td>
<td>66</td>
<td>12</td>
<td>(R) Frontoparietal</td>
<td>(L) Hemiparesis, impaired comprehension of figurative language</td>
<td>+</td>
</tr>
</tbody>
</table>

*Note. Age, years; postonset, months; VN, hemispatial visual neglect.*

“friend” in the example above) by stating (declarative) or seeking confirmation of (interrogative) the target information. Emotional readings of the stimuli were stimulated by providing an explicit context of the speaker’s reaction to the situational context prior to receiving the prompt (e.g., *You become very angry because you told Mary that morning not to include the teapot in the sale.*). Cue cards containing the target utterance for each trial served to visually reinforce the desired prosodic interpretation (bold italics were used to highlight focus position, appropriate punctuation facilitated modality distinctions, emotional target labels were written below each trial). Priming scenarios were recorded in a sound-attenuated chamber by a male speaker using a high-quality Sony tape recorder and a Sony ECM-909 directional microphone. The speaker was encouraged to produce the passages in a neutral, “reporting” tone that did not lend undue prominence to particular lexical items within the passage. The 48 scenarios were subsequently randomized for order of presentation and the tape was edited to reflect this...
random order. A 5-s interstimulus pause was inserted between passages to allow subjects
appropriate time to produce a response during the experiment.

Procedure

Subjects were tested individually in a quiet room in the subject’s home (RHD patients) or
in a speech laboratory (NC subjects). To maintain close attention to the content of the priming
scenarios, testing was completed during two 30-min sessions, half of the trials being presented
during each visit. Subjects were seated comfortably at a table with a directional microphone
(Sony ECM-909) placed 20 cm in front of their mouths and a small binder containing the
stimulus cards placed before them. Subjects were instructed to listen closely to each ‘‘story’’
(presented free-field using a Sony portable radio-cassette player) and then respond to the pas-
sage by reading the card placed in front of them. Subjects were asked to repeat sentences
when reading errors or dysfluencies occurred and were permitted to repeat their responses at
any time when not completely satisfied with their performance (such ‘‘corrections’’ were ob-
served highly infrequently). The final rendition produced for any given trial was always consid-
ered in subsequent analyses. Five practice trials preceding the experiment helped subjects
become acquainted with the elicitation procedure. Despite occasional comments on the repeti-
tious nature of the priming scenarios, little difficulty was experienced by either the NC or
RHD subjects in performing the task. All responses were recorded by a Sony (TCD-D3) digital
audiotape recorder for later analysis.

Acoustic Analyses

The 48 utterances generated for each speaker were digitized at a rate of 20 kHz (9-kHz
low-pass filter, 12-bit quantization) using the BLISS speech analysis system (Mertus, 1989)
and then acoustically analyzed using the BLISS waveform editor. Both short-term and long-
term measures of \( F_0 \) were computed to capture each group’s ability to project the various
combinations of emphatic, sentence modality, and emotional targets. Amplitude measures were
omitted owing to the minimal significance accorded to these cues by previous researchers
(Behrens, 1988; Bolinger, 1958; Morton & Jassem, 1965; Ross, Edmondson, Seibert, & Ho-
man, 1988; Turk & Sawusch, 1996; but cf. Sluijter, van Heuven, & Pacilly, 1997), whereas
temporal attributes of the stimuli were examined in detail for each speaker group but are
reported elsewhere (Pell, in press (b)). The following \( F_0 \) measures were derived for each utter-
ance.

1. \( \text{Mean } F_0 \). Following earlier studies of linguistic prosody (Cooper et al., 1985a; Eady &
Cooper, 1986; Eady et al., 1986), word-by-word transitions in \( F_0 \) were determined by measur-
ing mean \( F_0 \) for the stressed vowel of each “keyword” (i.e., content word) and within the
final 150 ms of the \( F_0 \) contour (Lieberman, 1967). Through visual inspection of the waveform,
five contiguous pulses were isolated at the center of the vowel (or within the terminal portion
of the utterance) by placing cursors at zero crossings and the inverted period was averaged
(Behrens, 1988; Ouellette & Baum, 1993). The mean \( \text{utterance } F_0 \) (i.e., the average of all
keyword \( F_0 \) values within a stimulus, including the utterance terminal portion) was also calcu-
lated and analyzed separately to exemplify gross differences in how different speakers set
mean \( F_0 \) for specific utterance types.

2. \( \text{Pitch accent}. \) The magnitude of \( F_0 \) change associated with content words with and with-
out focus was determined for sentence-initial and sentence-final keywords. As focused words
were always bisyllabic, “pitch accent” reflected the difference in mean \( F_0 \) between the initial
stressed vowel and the terminal portion of the second vowel (e.g., Ma-ry).
(3) $F_0$ variation. An indication of $F_0$ variation present in each utterance—perhaps one of the strongest perceptual cues to a speaker’s level of emotional activation (Ladd et al., 1985; Scherer, 1986; Williams & Stevens, 1972)—was characterized as the difference between the highest and the lowest $F_0$ values derived at all keyword points, including the utterance terminal portion.

**Statistical Analyses**

Acoustic measures were normalized prior to statistical inspection. To normalize for differences in $F_0$ range between subjects, mean $F_0$ values (keyword $F_0$, utterance $F_0$) were transformed to $z$ scores within speakers using the following formula: $\bar{F}_0^{\text{norm}} = (\bar{F}_0 - \bar{F}_0^{\text{mean}})/s$, where $\bar{F}_0$ is the observed $F_0$ value, $\bar{F}_0^{\text{mean}}$ is the mean $F_0$ across all utterances produced by the speaker and $s$ is the standard deviation (Colsher, Cooper, & Graff-Radford, 1987; Gandour et al., 1995). Measures of $F_0$ fluctuation (pitch accent, $F_0$ variation) were normalized for intersubject differences in mean $F_0$ by dividing the range values by the mean $F_0$ of the corresponding utterance for each speaker. Data obtained for the two test items were collapsed within each condition prior to statistical evaluation. All measures were examined statistically through analysis of variance (ANOVA) in a full factorial design. Group membership (NC, RHD) served as the between-subjects factor for each analysis and repeated measures constituted a subset of the following: MODALITY (declarative, interrogative), FOCUS (initial, final, none), EMOTION (neutral, sad, happy, angry), and KEYWORD (K1–K4; review Table 2). Statistically significant effects were explored posthoc using Tukey’s HSD procedure ($p < .05$), wherever appropriate.

**Perceptual Analyses**

A subset of utterances elicited from each NC and RHD subject were rated by 10 young normal listeners to gauge the perceptibility of emphasis and emotion in their speech. Declarative readings of one of the two stimulus items (“Mary sold the teapot”) were presented to subjects individually over headphones in two separate sessions. On one occasion, listeners were presented those stimuli elicited from NC and RHD subjects in a sad, happy, or angry tone (3 emotions × 3 focus positions × 10 subjects × 2 groups = 180 trials) and were asked to choose the emotional tone that “most closely fit” that used by the speaker for each exemplar (a push-button response was recorded). Perceptual raters were offered four possible interpretations in judging the emotional prosody including a “neutral” designation (although note that none of the perceptual trials had been elicited in a “neutral” mode). A second session presented a highly overlapping subset of the stimuli in which emphasis had been elicited in either sentence-initial or sentence-final position (2 focus positions × 4 emotions × 10 subjects × 2 groups = 160 trials). For the focus perception task, the 10 listeners were asked to identify the position in each utterance where they felt emphasis was present, if any (alternatives: initial, final, none). For both perception tasks, stimuli elicited from the RHD and NC subjects were fully randomized within the same experiment and raters were unaware of group distinctions within the data. Half of the perceptual raters received the emotion perception task first and the other half performed the focus perception task first.

**RESULTS**

Statistically significant findings are reported independently by acoustic measure. Given the number of independent factors considered in the present design, the text is confined to a description of those significant effects
<table>
<thead>
<tr>
<th>Group</th>
<th>Modality</th>
<th>Focus</th>
<th>K1</th>
<th>K2</th>
<th>K3</th>
<th>K4</th>
<th>Neutral</th>
<th>Sad</th>
<th>Happy</th>
<th>Angry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mary</td>
<td>sold</td>
<td>tea-</td>
<td>pot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.42</td>
<td>-1.71</td>
<td>-1.81</td>
<td>-2.59</td>
<td>-0.47</td>
<td>-1.60</td>
<td>-1.67</td>
<td>-2.07</td>
</tr>
<tr>
<td>NC</td>
<td>(.</td>
<td>K1</td>
<td>0.58</td>
<td>0.96</td>
<td>0.85</td>
<td>3.13</td>
<td>-0.68</td>
<td>0.30</td>
<td>0.49</td>
<td>2.68</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>K3</td>
<td>-0.72</td>
<td>-0.73</td>
<td>1.80</td>
<td>-2.22</td>
<td>-0.42</td>
<td>-0.82</td>
<td>-0.21</td>
<td>-2.13</td>
</tr>
<tr>
<td></td>
<td>(??)</td>
<td></td>
<td>-0.44</td>
<td>-0.44</td>
<td>0.42</td>
<td>2.98</td>
<td>-0.73</td>
<td>-0.87</td>
<td>-0.16</td>
<td>1.70</td>
</tr>
<tr>
<td>RHD</td>
<td>None</td>
<td>K1</td>
<td>0.58</td>
<td>-1.71</td>
<td>-1.80</td>
<td>-2.46</td>
<td>-0.01</td>
<td>-1.50</td>
<td>-1.73</td>
<td>-2.30</td>
</tr>
<tr>
<td></td>
<td>(??)</td>
<td></td>
<td>-0.37</td>
<td>-0.61</td>
<td>1.47</td>
<td>-2.18</td>
<td>-0.21</td>
<td>-0.37</td>
<td>0.38</td>
<td>-1.50</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>K3</td>
<td>-0.28</td>
<td>-0.12</td>
<td>0.82</td>
<td>1.89</td>
<td>-0.21</td>
<td>-0.12</td>
<td>0.73</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.06</td>
<td>0.42</td>
<td>1.40</td>
<td>2.23</td>
<td>-0.14</td>
<td>-0.42</td>
<td>0.50</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Note: (. Declarative utterances; (?) interrogative utterances; focus, position of emphasis within the utterance; K, keyword.
deemed of greatest theoretical relevance. A table of all statistically significant effects produced by each analysis is furnished in the Appendix.

(1) Mean $F_0$

Mean (normalized) $F_0$ values derived at the four keyword points of interest are supplied in Table 2 as a function of sentence modality, focus position, and emotional context. To underscore potential group differences in the modulation of short-term linguistic parameters as a function of affective speaking mode, the keyword $F_0$ data were first analyzed using four GROUP \times MODALITY \times FOCUS \times KEYWORD ANOVAs, performed independently by emotional context. Results obtained from these analyses pointed to a number of important differences in how RHD and NC speakers used $F_0$ to mark focus and modality distinctions as a function of some (but not all) of the emotional contexts.

For utterances produced in a sad tone, the ANOVA yielded significant interactive effects of GROUP \times MODALITY \times FOCUS \times KEYWORD [$F(2, 36) = 5.73, p < .01$] and GROUP \times MODALITY \times KEYWORD [$F(3, 54) = 3.95, p < .05$]. Tukey’s pairwise comparisons performed on the first interaction indicated that focus distribution did not affect the overall level at which normal speakers set mean $F_0$ for sad utterances, but did influence RHD speakers somewhat in this condition (utterances containing terminal focus displayed a significantly higher $F_0$ than those containing initial or no focus, overall). Consequently, RHD patients tended to produce utterances with sentence-final focus with a significantly higher (overall) $F_0$ level than that typical of normal speakers when speaking in a sad tone. Exploration of the interaction among GROUP, MODALITY, and KEYWORD revealed that when marking differences in sentence modality, RHD patients tended to elevate mean $F_0$ at keyword positions preceding the utterance terminal (K3) in interrogative utterances when compared to declarative utterances, contrary to the normal “flat” pattern of all sad utterances in this portion of the contour. Although both groups robustly marked the declarative/interrogative distinction at the utterance terminal (K4) through the expected fall/rise of the $F_0$ contour, it is noteworthy that RHD speakers marked interrogation with a significantly lower $F_0$ rise than did the NC subjects in the sad condition.

Analysis of stimuli elicited in a happy tone produced a significant main effect for GROUP [$F(1, 18) = 7.96, p = .01$] and a four-way interaction of GROUP \times MODALITY \times FOCUS \times KEYWORD [$F(6, 108) = 2.45, p < .05$], illustrated in Fig. 1 by group and sentence modality. Inspection of the interaction suggested that when producing declarative utterances in a happy tone, RHD patients did not significantly elevate the $F_0$ of sentence-initial words (K1) in the focused condition relative to the unfocused condition, contrary to normal speakers. Furthermore, RHD patients exhibited a nonsignificant rise in $F_0$ when marking focus on sentence-final items (i.e.,
FIG. 1. Effect of sentence modality, focus position, and keyword location on utterances produced in a happy tone by normal and RHD speakers.

between K2 and K3), contrary to the normal pattern. When producing interrogative contours, RHD patients exhibited a significantly lower terminal F0 (K4) than NC subjects when speaking in a happy tone for two of the three focus conditions (no focus and sentence-final focus). Group differences in terminal measures appeared to emerge as a result of the RHD patients’ failure to produce a significant rise in F0 in this position for any of the three focus conditions, as measured between K3 and K4 (review Fig. 1).

The keyword analysis of angry utterances yielded a significant interaction of GROUP × FOCUS \(F(2, 36) = 4.08, p < .05\). Similar to results obtained for sad stimuli, RHD patients demonstrated a unique sensitivity to the presence/location of focus in regulating the overall mean F0 level of angry utterances. Although normal speakers produced angry utterances varying in focus location with a uniform F0 level, the RHD speakers did not display this pattern (utterances with initial focus displayed a lower F0 than those with final or no focus for the RHD patients). No group differences emerged in the keyword F0 data when angry utterances matched in focus location were compared.

Unlike the three contexts in which emotional features were present, the ANOVA performed on utterances elicited in a neutral (i.e., nonaffective) tone did not lead to significant main or interactive effects with the GROUP factor \((p > .05)\). The absence of group differences in this condition at any
of the four keyword locations points to a high degree of correspondence in
the ability of the NC and RHD speakers to mark emphasis and modality
distinctions via F0 when the “load” of particular combinations of prosodic
variables was (presumably) lowest (or when particular properties of intona-
tion contours may have been limited—see Discussion).

A direct comparison of how each group encoded linguistic– and emo-
tional–prosodic representations in tandem was achieved through a 2
(GROUP) × 2 (MODALITY) × 3 (FOCUS) × 4 (EMOTION) ANOVA
performed on the mean utterance F0 values (i.e., the average of the four key-
word measures). Table 3 summarizes these data by emotion type, sentence
modality, and focus position (following normalization). A significant three-
way interaction of GROUP × EMOTION × FOCUS [F(6, 108) = 2.79,
p < .05], displayed in Fig. 2, indicated that RHD patients produced happy
prosody with a significantly lower mean F0 than NC subjects in two of the
three focus conditions (utterances with sentence-final and no focus). In con-
trast, no group differences were observed in the production of neutral, sad,
or angry prosody as a function of emphasis placement.

Further examination of the mean utterance F0 data indicated that the two
groups varied significantly in the ability to differentiate the four emotional
tones via mean F0. Although normal subjects always expressed happy pros-
ody with a significantly elevated mean F0 relative to that of neutral and sad
prosody, RHD patients elevated the F0 of happy sentences only when com-
pared to sad sentences and then only in one context (when sentence-initial
focus was present).1 These findings suggest that the RHD patients may have
been disturbed in setting F0 at a level appropriate to happy prosody when
producing the stimuli, leading to a reduction in the number of distinctions
among the four emotions with respect to mean F0.

Finally, important patterns emerged in the mean F0 data which describe
the interplay of linguistic modality, focus, sentence position, and emotion
independent of group affiliation. Notably, a significant interaction of
MODALITY × FOCUS × KEYWORD was produced by each of the four
keyword analyses [FNEUTRAL,6, 108] = 19.76; FSAD,6, 108] = 5.35; FHAPPY,6,
108] = 11.60; FANGRY,6, 108] = 8.01; p < .001 in all cases], depicted in
Fig. 3 (independently by emotional context). The emergence of this interac-
tion corroborates and extends earlier work on nonaffective (neutral) speech
in normal individuals (Cooper et al., 1985a; Eady & Cooper, 1986) and has
been treated in detail by Pell (1997a). Nonetheless, it is constructive to reem-
phasize the complex (yet systematic) relationship between various aspects
of prosodic structure in the acoustic signal; in particular, the extent to which

1 The NC subjects also produced significantly higher mean F0 values for the following emo-
tions: happy relative to angry (final focus), angry relative to sad (final and no focus), and
neutral relative to sad (final focus). The RHD patients also produced angry prosody with
significantly higher F0 than sad prosody in one context (utterances without focus).
TABLE 3
Mean Normalized F₀ and F₀ Range of the Utterances Produced by the NC and RHD Speakers, as a Function of Sentence Modality, Focus Position, and Emotional Context

<table>
<thead>
<tr>
<th>Group</th>
<th>Modality</th>
<th>Focus</th>
<th>Mean F₀</th>
<th>F₀ range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Neutral</td>
<td>Sad</td>
</tr>
<tr>
<td>NC</td>
<td>(·)</td>
<td>K1</td>
<td>-1.42</td>
<td>-1.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K3</td>
<td>-0.46</td>
<td>-0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>None</td>
<td>-0.95</td>
<td>-1.04</td>
</tr>
<tr>
<td></td>
<td>(?)</td>
<td>K1</td>
<td>1.09</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K3</td>
<td>0.63</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>None</td>
<td>0.55</td>
<td>0.32</td>
</tr>
<tr>
<td>RHD</td>
<td>(·)</td>
<td>K1</td>
<td>-1.30</td>
<td>-1.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K3</td>
<td>-0.42</td>
<td>-0.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>None</td>
<td>-0.87</td>
<td>-1.06</td>
</tr>
<tr>
<td></td>
<td>(?)</td>
<td>K1</td>
<td>0.97</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K3</td>
<td>0.58</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>None</td>
<td>1.04</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Note. (·) Declarative utterances; (?) interrogative utterances; K1, emphasis on sentence-initial word; K3, emphasis on sentence-final word.
interrogative contours impose limitations on how F0 may be used to mark emphasis across focus conditions is highlighted by Fig. 3, relative to the same forms encoded in conjunction with declarative contours.

(2) *Pitch Accent*

Table 4 presents (normalized) data describing the magnitude of F0 fluctuation associated with sentence-initial (K1) and sentence-final (K3) keywords for the two diagnostic groups, as a function of sentence modality, focus placement, and emotional context.

A five-way (GROUP × EMOTION × MODALITY × FOCUS × KEYWORD) ANOVA performed on these data indicated that RHD patients produced significantly less F0 variation than normal speakers, overall, on content words important in signaling linguistic prominence in the current stimulus set [GROUP main effect: $F(1, 18) = 5.97, p < .05$]. No interactions
between the group factor and the four independent prosodic variables were observed, pointing to the (otherwise) high degree of conformity in how the two groups encoded emphatic stress cues across stimulus types. In particular, the absence of a GROUP × FOCUS × KEYWORD interaction from this analysis suggests that the restriction in F₀ modulation exhibited by the RHD patients was likely of a generalized nature and did not represent a specific defect in producing emphatic stress cues per se.

A significant four-way interaction of EMOTION × MODALITY × FOCUS × KEYWORD also emerged \( F(6, 108) = 2.38, p < .05 \), presented in Fig. 4 (independently by sentence modality). Summarized briefly, words in both sentence-initial (K1) and sentence-final (K3) positions were produced with significantly greater F₀ variation when focused than when

---

**FIG. 3.** Influence of sentence modality, focus position, and keyword location on mean F₀ for utterances spoken in the four emotional tones.
<table>
<thead>
<tr>
<th>Group</th>
<th>Modality</th>
<th>Focus</th>
<th>Neutral</th>
<th>Sad</th>
<th>Happy</th>
<th>Angry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>K1 Ma-ry</td>
<td>K3 tea-pot</td>
<td>K1 Ma-ry</td>
<td>K3 tea-pot</td>
<td>K1 Ma-ry</td>
</tr>
<tr>
<td>NC</td>
<td>(.)</td>
<td>0.43</td>
<td>0.19</td>
<td>0.26</td>
<td>0.16</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>K3</td>
<td>0.14</td>
<td>0.84</td>
<td>0.17</td>
<td>0.42</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>0.10</td>
<td>0.46</td>
<td>0.17</td>
<td>0.37</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(?)</td>
<td>0.08</td>
<td>0.59</td>
<td>0.25</td>
<td>0.38</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>0.13</td>
<td>0.65</td>
<td>0.12</td>
<td>0.46</td>
<td>0.21</td>
</tr>
<tr>
<td>RHD</td>
<td>(.)</td>
<td>0.40</td>
<td>0.15</td>
<td>0.31</td>
<td>0.14</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>K3</td>
<td>0.11</td>
<td>0.64</td>
<td>0.13</td>
<td>0.39</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>0.18</td>
<td>0.31</td>
<td>0.11</td>
<td>0.24</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>K1</td>
<td>0.24</td>
<td>0.35</td>
<td>0.14</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(?)</td>
<td>0.10</td>
<td>0.53</td>
<td>0.12</td>
<td>0.34</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>0.12</td>
<td>0.41</td>
<td>0.12</td>
<td>0.33</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*Note: (.) Declarative utterances; (?) interrogative utterances; focus, position of emphasis within the utterance; K, keyword.*
FIG. 4. Effect of sentence modality, focus position, and emotional tone on the amount of $F_0$ change produced on sentence-initial (K1) and sentence-final (K3) keywords. (a) Declarative; (b) interrogative.

unfocused, across stimulus types. Differences in sentence modality had limited (albeit some) effect on local alterations in $F_0$, whereas the impact of emotional features on the production of pitch accents was evident solely for sentence-final keywords; in this position, sad sentences displayed consistently less variation than happy sentences and, to a somewhat lesser extent, angry and neutral sentences (review Fig. 4). These findings suggest that the magnitude of $F_0$ accents underlying linguistic focus are influenced by the emotional mode of the speaker, particularly toward the end of the utterance.

2 Exceptionally, sentences spoken in a sad tone often displayed nonsignificant differences in $F_0$ fluctuation in focused relative to unfocused tokens, at both sentence positions.
with larger $F_0$ accents being associated with emotions characterized by greater long-term variation in $F_0$ (e.g., happy, angry—see below).

(3) $F_0$ Variation

Generalized variability in $F_0$ may be an especially salient marker of emotional states (e.g., Bolinger, 1978; Ladd et al., 1985; Scherer, 1986). Normalized $F_0$ variation values produced by each of the diagnostic groups are provided in Table 3 (with mean utterance $F_0$ values) as a function of emotional mode, sentence modality, and focus location.

A four-way (GROUP $\times$ EMOTION $\times$ MODALITY $\times$ FOCUS) ANOVA performed on the $F_0$ variability data produced a significant main effect for GROUP [$F(1, 18) = 5.33, p < .05$] but did not yield any significant group interactions ($p > .05$). Consistent with results obtained for the pitch accent data, the group main effect was explained by a generalized decrease in the amount of $F_0$ variation produced by the RHD patients across stimulus types when compared to the NC speakers. Figure 5 highlights the extent of $F_0$ variation produced by each of the 10 normal and each of the 10 RHD speakers when expressing the four emotions. Although a reduction in $F_0$ variability was not shown to be tied to a given linguistic or affective speaking mode when examined statistically, Fig. 5 suggests a trend whereby “happy” (and to a lesser extent “angry”) may have exhibited the least normal distribution in the RHD patients’ speech due to this group’s restricted $F_0$ variation. In contrast, the representation of neutral and sad contours appeared to be less affected by the (generalized) lack of $F_0$ variation in the RHD patients’ speech based on qualitative analysis of the data.

Analysis of the $F_0$ variation data also uncovered a significant EMOTION $\times$ MODALITY interaction [$F(3, 54) = 3.45, p < .05$], represented in Fig. 6. For both declarative and interrogative stimuli, sad prosody was shown to be distinct in its lack of $F_0$ variation relative to the other three emotions. Happy prosody was characterized by significantly more $F_0$ variability than angry prosody for both modality types and than neutral prosody for declarative sentences. Comparison of range values in matched declarative and interrogative sentences suggested that only happy prosody differed as a function of sentence modality, $F_0$ variation proving significantly constrained when spoken in the interrogative mode than in the declarative mode.

Perceptual Ratings

The success of each of the 20 speakers in transmitting emphatic and emotional distinctions to normal listeners may be viewed in Table 5. Both RHD (68.1%) and NC (78.6%) speakers signaled emphasis to listeners at a level well exceeding chance (33.3%). A two-way (GROUP $\times$ FOCUS) ANOVA performed on the emphasis ratings indicated that RHD speakers were significantly less reliable at conveying linguistic prominence than normal speak-
FIG. 5. Amount of $F_0$ variation produced by normal and RHD speakers (shown independently by subject) for each of the four emotional tones.

ers overall [GROUP main effect: $F(1, 18) = 14.46, p = .001$], but revealed a disparity in group performance as a function of the *sentential position* in which emphatic stress cues were encoded [GROUP $\times$ FOCUS interaction: $F(1, 18) = 32.43, p < .001$]. Normal speakers were equally capable of communicating linguistic focus in sentence-initial (80%) and sentence-final (77%) position, whereas RHD speakers displayed inferior encoding of sentence-final (56%) relative to sentence-initial (80%) emphasis. The performance of the two groups differed significantly only for sentence-final targets, for which RHD speakers were judged to be far less accurate than normal speakers (see Fig. 7a).

Similarly, the success of the NC and RHD groups in encoding emotional
FIG. 6. Effect of sentence modality on the amount of $F_0$ variation produced for each of the four emotional tones.

TABLE 5
Percentage of Correct Identifications of Focus Position and Emotional Tone for Each Normal (NC) and Right Hemisphere-Damaged (RHD) Speaker (Averaged across the 10 Listeners)

<table>
<thead>
<tr>
<th>Focus perception (chance level = 33%)</th>
<th>Emotion perception (chance level = 25%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject</strong></td>
<td><strong>NC</strong></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>88</td>
</tr>
<tr>
<td>3</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>98</td>
</tr>
<tr>
<td>6</td>
<td>90</td>
</tr>
<tr>
<td>7</td>
<td>65</td>
</tr>
<tr>
<td>8</td>
<td>91</td>
</tr>
<tr>
<td>9</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>78.6</strong></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td><strong>±13.7</strong></td>
</tr>
</tbody>
</table>

*Note.* Values are expressed as an average rating across the two focus positions (initial, final) or three emotional tones (sad, happy, angry) for each subject. Calculation of chance performance considered the total number of response alternatives offered to raters on each task, including a “no focus” (focus perception) or “neutral” (emotion perception) designation.
distinctions within the same sentences was well above chance level when viewed overall (NC = 67.2%; RHD = 49.5%; chance performance = 25%). Results of a two-factor (GROUP × EMOTION) ANOVA indicated that normal speakers again surpassed RHD speakers on the emotion perception task overall [GROUP main effect: $F(1, 18) = 77.85, p < .001$], but that this effect only applied to specific emotions [GROUP × EMOTION interaction: $F(2, 36) = 19.09, p < .001$]. Specifically, normal speakers were more proficient in transmitting happy prosody (78%) to listeners than either angry (64%) or sad (61%) prosody (which did not differ), whereas RHD speakers exhibited the opposite pattern: accuracy for happy prosody (37%) was sig-
significantly lower than that for both angry (58%) and sad (56%) prosody (examine Fig. 7b). (See Gandour et al., 1995, for a similar reversal in the identification of happy prosody when produced by NC and RHD speakers.) Thus, when the ability to express each emotion was examined independently, significant group differences were noted exclusively for happy prosody (NC = 78%, RHD = 37%). The perceptual findings appear to confirm signs in the acoustic data (especially for mean \( F_0 \), \( F_0 \) variation) that the production of ‘‘happy’’ prosody may have been most aberrant in the RHD patients’ speech.

**DISCUSSION**

There is little debate that fundamental frequency serves as the primary vehicle by which many intonational meanings are specified in English, whether those features be linguistically defined or more biologically driven. Delineating the capacity of RHD native English speakers to modulate \( F_0 \) when multiple prosodic distinctions—both linguistic and emotional in nature and those encompassing both ‘‘local’’ and ‘‘global’’ acoustic domains—were present concurrently in experimental stimuli represents an initial attempt to understand the interactive influence of these prosodic dimensions on right hemisphere function. These data may illuminate elements of prosodic structure sensitive to right hemisphere compromise in speech production, shedding light on previous hypotheses of prosody lateralization (Behrens, 1989; Ross, 1981; Van Lancker & Sidtis, 1992). The results may further underscore important attributes of prosody production in normal speakers.

The cues to linguistic prominence, although specified through local alterations in the \( F_0 \) and duration of a focused constituent (Behrens, 1988; Cooper, Soares, & Reagan, 1985b), operate as a relative measure of acoustic change occurring across sentential units (e.g., Bolinger, 1961; Gussenhoven, Repp, Rietveld, Rump, & Terken, 1997; Liberman & Pierrehumbert, 1984). The success of RHD patients in implementing the \( F_0 \) correlates of focus was investigated at strategic points in their utterances to monitor for irregularities in both the direction and relative magnitude of \( F_0 \) change produced. Despite current evidence that the \( F_0 \) correlates of linguistic focus vary as an index of sentential position and linguistic modality (also Eady & Cooper, 1986), the production of emphasis by the RHD patients was shown to pattern in a relatively similar manner to that of the normal subjects in many important respects. When producing declarative utterances, both groups exhibited a large postfocus drop in \( F_0 \) when sentence-initial words were emphasized and a significant rise in \( F_0 \) when sentence-final words were emphasized (Cooper et al., 1985a; Eady & Cooper, 1986). When producing interrogative utterances, both groups conveyed sentence-initial focus via a nonsignificant rise in mean \( F_0 \) on subsequent words (Eady & Cooper, 1986). Furthermore, RHD speakers displayed the normal tendency to produce words receiving focus with greater \( F_0 \) variation than matched words without focus.
In fact, the RHD patients’ qualitatively normal use of $F_0$ to signal focus across a wide variety of stimulus types (a corpus of nearly 250 utterances was scrutinized for each group) is viewed as strong evidence that knowledge of emphatic stress production was largely spared in the current RHD sample (Behrens, 1988; Ouellette & Baum, 1993). This outcome argues in favor of minimal right hemisphere involvement in the processes that encode linguistic stress features at the word level, as postulated on previous occasions (Behrens, 1988, 1989; Emmorey, 1987). Such a conclusion also fits with recently published data demonstrating a spared capacity to perceive emphatic stress by RHD (but not LHD) patients (Baum, 1998; Pell, 1998).

It is noteworthy that previous investigations reporting a defect in the production of emphatic stress by RHD patients are conspicuous in their analysis of relatively few experimental stimuli and their strict reliance on auditory-perceptual ratings rather than a combination of acoustic and perceptual measures (Brådvik et al., 1991; Bryan, 1989; Weintraub et al., 1981). As suggested by Behrens’ (1988) data, the production of ‘‘normal’’ acoustic parameters to emphasis by RHD patients may not always be reflected in emphasis cues that are of normal perceptual saliency to the listener. Indeed, the current RHD sample was significantly inferior to normal speakers in conveying emphatic stress position to listeners based on the perceptual data, despite evidence that ‘‘pitch accents’’ produced by each subject group were similar in overall shape and direction. This apparent discrepancy may be related to the observation that pitch accents, although appropriate when categorical properties of the stimuli were accessed, were significantly reduced in overall magnitude when elicited from the RHD speakers. Thus, not all of the acoustic cues to focus (i.e., continuous properties) were entirely normal in the present clinical sample. As such, previous claims of ‘‘reduced’’ or impaired emphasis production in RHD patients derived from auditory–perceptual ratings are not necessarily discordant with the idea that RHD patients retain knowledge of how linguistic focus is encoded in the speech signal, but are less reliable in implementing all of the normal cues necessary to perceive emphasis when judged from the listener’s standpoint (Brådvik et al., 1991; Bryan, 1989; Weintraub et al., 1981).

In the production of sentence modality distinctions, the acoustic data (not surprisingly) revealed that the direction of $F_0$ fluctuation observed in the final 150 ms of the utterance reliably distinguished the two illocutionary modes investigated, interrogative contours exhibiting a marked rise in terminal $F_0$ versus a terminal decline for declarative contours (Behrens, 1989; Eady & Cooper, 1986; Ohala, 1984; Studdert-Kennedy & Hadding, 1973). Sentence modality markers were also prevalent at points preceding the terminal; most notably, in the case of sentences containing sentence-initial focus, the direction of the postfocus $F_0$ excursion corresponded to the marked (interrogative) or unmarked (declarative) direction of the $F_0$ terminal (Eady & Cooper, 1986). In general, declarative and interrogative contours elicited from the
RHD patients reported herein were reflective of normal patterns of cue usage, with some exceptions. Although RHD speakers demonstrated consistent use of F0 in differentiating declaratives from interrogative at the terminal portion of the utterance (interrogatives were always significantly higher in F0 than declaratives), the RHD patients’ ability to encode interrogative contours in conjunction with emotional prosody appeared to be aberrant at times. For example, RHD patients marked interrogation with a significantly smaller terminal F0 rise than NC subjects when speaking in a sad tone and failed altogether to produce a significant terminal rise in F0 when speaking in a happy tone (see Klouda et al., 1988, for similar trends). Despite such irregularities when producing interrogative stimuli, measures of terminal declination underlying declarative utterances did not significantly differ for the two groups as a function of emotional content.

Thus, in keeping with the emphatic stress data, the RHD patients demonstrated intact knowledge of linguistic (categorical) features of intonation contours over a vast number of utterances through the successful modulation of terminal F0 features to express sentence modality distinctions. However, the concurrent demands of modulating F0 to signal interrogation (the “marked” intonational feature according to Lieberman, 1967) and to signal emotional attributes highly reliant on F0 parameters (“happy” and “sad” are normally expressed with maximal and minimal F0 variation within the current stimulus set) may have proven too great a burden for the RHD patients, yielding abnormalities in their speech. This result suggests that the prosodic “load” specified by individual acoustic features such as mean F0 may be a factor in the expressive performance of RHD patients. Evidence that linguistic and emotional sentence prosody may exert an interactive influence on the expressive performance of RHD individuals has not previously been documented for native speakers of a nontone language such as English (but see Edmondson et al., 1987; Gandour et al., 1995, for comments on the impact of phonemic tone on affective signaling in RHD speakers of Taiwanese or Thai).

The ability of RHD patients to produce emotional–prosodic contrasts was determined primarily through long-term acoustic measures averaged across the utterance as a whole (Fairbanks & Pronovost, 1939; Huttar, 1968; Williams & Stevens, 1972). Generally, it was shown that the four affect types elicited in the present study separated well as a function of mean F0 and F0 variation. When compared to neutral prosody, happy prosody tended to exhibit an elevated mean F0 and larger F0 dispersion overall. Sad prosody displayed a similar mean F0 but diminished F0 range when compared to neutral prosody. Finally, angry prosody resembled neutral prosody in F0 variation but exhibited a significantly elevated mean F0. (Differences in speech rate further differentiated this stimulus set—see Pell, in press (b)). Thus, when compared to emotionally inflected utterances, neutral utterances were characterized by a relatively low mean F0 and an intermediate amount of F0 variation. This acoustic profile conforms in a general manner to previous descrip-
tions in the literature on emotional communication (Fairbanks & Hoaglin, 1941; Fairbanks & Pronovost, 1939; Frick, 1985; Pell & Baum, 1997b; Scherer, 1986; Williams & Stevens, 1972).

The RHD speakers’ proficiency in modulating the acoustic underpinnings of discrete emotions was of particular interest, given repeated claims that this skill is specifically dependent on intact right hemisphere function (Borod et al., 1990; Edmondson et al., 1987; Gandour et al., 1995; Ross, 1981; Shapiro & Danly, 1985). For mean utterance F$_0$, similar qualitative tendencies in how the two groups marked distinctions across emotional conditions could be inferred from the data (review Fig. 5) (Baum & Pell, 1997; Gandour et al., 1995). Despite such trends, the RHD patients’ use of mean F$_0$ departed significantly from normalcy when producing happy prosody, the patient group demonstrating a failure to elevate F$_0$ to the extent typical of normal (happy) speakers. Little evidence of irregularities in the programming of mean F$_0$ emerged for sad, angry, or neutral stimuli between the two groups. Overall, these data point to a subtle deficit in the ability of RHD patients to regulate mean F$_0$ in the communication of discrete affective messages (Colsher et al., 1987; Dykstra et al., 1995; Edmondson et al., 1987; Gandour et al., 1995; Shapiro & Danly, 1985).

Analysis of long-term measures of F$_0$ variation demonstrated a more pervasive impairment than that observed for mean F$_0$. RHD speakers displayed a significant restriction in intonational variation across stimulus types when compared to normal speakers (group main effect). An attenuation of F$_0$ variation constitutes one of the most frequent acoustic abnormalities attributed to RHD speakers in the literature on prosody (Baum & Pell, 1997; Behrens, 1989; Blonder et al., 1995; Colsher et al., 1987; Edmondson et al., 1987; Kent & Rosenbek, 1982; Klouda et al., 1988; Shapiro & Danly, 1985) and obtains further support herein. Moreover, inspection of each group’s productions across emotional conditions suggested a trend wherein happy prosody (which was characterized by extensive F$_0$ variation) diverged to a greater extent from normalcy than the other three emotions due to the lack of F$_0$ variation in the RHD patients’ speech. This emotion-specific trend, which did not reach significance based on the present analysis of F$_0$ variation, appeared to be corroborated by the perceptual ratings, where RHD speakers demonstrated a specific lack of facility in communicating happy prosody to listeners, the exact opposite pattern to that of normal speakers.

The RHD patients’ selective difficulty in encoding happy prosody may reflect (in part) a right hemisphere propensity to modulate acoustic properties which were most typical of happy contours within the present stimulus set. Specifically, happy contours were characterized by extreme departures from

---

3 An emotion-specific attenuation in the amount of F$_0$ variation employed to express happy prosody was significant for the same RHD sample when longer sentences were elicited using the same paradigm (Pell, 1997a).
neutral contours (i.e., the “unmarked” affective mode, Bolinger, 1986; Gandour et al., 1995; Ladd, 1980) in terms of both mean $F_0$ and $F_0$ variation; these differences may have rendered happy stimuli most difficult for the RHD speakers to implement in a consistent manner. This possibility relies on the notion that emotional messages are expressed along a continuum of acoustic change (Ohala, 1984; Scherer, 1986) and that increased physiological activation (indicative of “happy’’) tends to be conveyed through elevation of the “topline” or $F_0$ maximum within the intonation contour, relative to a neutral baseline (Cohen, Collier, & ’t Hart, 1982; Protopapas & Lieberman, 1997). Based on patterns observed in the present data, RHD speakers may have retained the capacity to implement local (emphatic stress) and global (sentence modality) categorical changes in intonation contours, but their control of continuous aspects of prosodic phenomena may have been deficient (Blonder et al., 1995). Such a defect, although most prevalent in the encoding of emotional messages (such as happy), was also reflected in the RHD patients’ inability to encode continuous properties of certain linguistic forms, such as the magnitude of $F_0$ change underlying pitch accents or the extent of the terminal rise in interrogative contours (Blonder et al., 1995; Klouda et al., 1988). Collectively, these largely quantitative irregularities may be indicative of a mechanical or phonetic deficit for prosody following right hemisphere insult, one which becomes most prevalent when encoding the rise portion of $F_0$ contours (Blonder et al., 1995; Ryalls, Joanette, & Feldman, 1987). These preliminary conclusions await further research to test their validity.

It is important to highlight that the global effect of the RHD patients’ imprecision in regulating continuous aspects of prosody was one of reduced acoustic distinctiveness across stimulus types, or a tendency for utterances varying in emotional and emphatic content to “converge” acoustically (Behrens, 1988; Blonder et al., 1995; Klouda et al., 1988; Pell, 1997a; for an example, review Fig. 1). A tendency to produce fewer or smaller distinctions in the prosodic line of an utterance may contribute strongly to the subjective impression of reduced intonation or “flat affect” in many RHD speakers (Baum & Pell, 1997; Borod et al., 1990; Ross, 1981). Indeed, the likelihood that many of the RHD patients described herein would be perceived as emotionally “flat” in their speech was suggested by a recent follow-up study conducted on the present data; as reported by Pell (in press (a)), low ratings on the emotion perception task were strongly associated with a failure to adequately distinguish the three emotional tones in acoustic terms for at least half of the current RHD sample.

If indeed gradient aspects of prosody pose specific difficulties for RHD patients in speech production, it is not surprising that deficits in expressing emotional distinctions (which are related to the speaker’s level of activation and therefore inherently “continuous” in nature) have most frequently been
ascribed to this clinical group (Borod et al., 1990; Edmondson et al., 1987; Gandour et al., 1995; Ross, 1981; Shapiro & Danly, 1985). In a similar vein, it is not unexpected that few studies have reported substantive differences in the capacity of RHD and normal subjects to produce linguistic intonation contours specified by categorical modifications in acoustic content, such as the declarative/interrogative distinction (Baum & Pell, 1997; Behrens, 1989; Blonder et al., 1995; Cooper et al., 1984; Gandour et al., 1995; Lebrun, Lessinnes, De Vresse, & Leleux, 1985; Ryalls et al., 1987; but cf. Brådvik et al., 1991; Shapiro & Danly, 1985; Weintraub et al., 1981). This pattern of performance may be predicted by the greater importance of graded acoustic phenomena in signaling affective versus “grammatical” prosodic meanings in verbal behavior. Remaining conflicts in the literature on linguistic intonation may reflect, in part, idiosyncrasies in cue use within experimentally defined speaking modes and the likelihood that “neutral” utterances are seldom completely devoid of speaker affect and thus, emotional highlighting via continuous acoustic information.

In interpreting the present data, it is imperative to note that matched patients with unilateral left hemisphere insult and aphasia were not examined due to the excessive demands of the elicitation procedure and the complexity of the verbal stimuli. The absence of such a control group allows for the present contentions regarding the right hemisphere’s role in modulating graded aspects of prosodic contours, but constrains the ability to draw implications about interhemispheric processes underlying prosody encoding. Examination of the ability to modulate intonation in subjects with analogous lesions in left and right hemisphere regions of the brain will be necessary before firm conclusions about the right hemisphere’s dominance for continuous aspects of prosody can be established. As well, a more detailed examination of individual performance characteristics with respect to intrahemispheric lesion site would benefit future undertakings that consider the effects of unilateral brain damage on the continuous and categorical attributes of speech prosody.
### APPENDIX: Significant ANOVA Results

**Main Effects**

<table>
<thead>
<tr>
<th></th>
<th>Interactions with group</th>
<th>Interactions without group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mean F&lt;sub&gt;0&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>M: F(1, 18) = 104.98, p &lt; .001</td>
<td>GxF: F(2, 36) = 5.73, p &lt; .01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GxMxK: F(3, 54) = 3.95, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MxK: F(3, 54) = 57.33, p &lt; .001</td>
</tr>
<tr>
<td>Happy</td>
<td>G: F(1, 18) = 7.96, p = .01</td>
<td>GxMxK: F(3, 54) = 5.79, p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>M: F(1, 18) = 68.77, p &lt; .001</td>
<td>GxMxKx: F(6, 108) = 2.45, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>K: F(3, 54) = 6.67, p = .001</td>
<td>FxK: F(6, 108) = 4.05, p = .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MxFxK: F(6, 108) = 15.23, p &lt; .001</td>
</tr>
<tr>
<td>Angry</td>
<td>M: F(1, 18) = 73.23, p &lt; .001</td>
<td>GxF: F(2, 36) = 4.08, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>F: F(2, 36) = 12.05, p &lt; .001</td>
<td>MxF: F(6, 108) = 67.95, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>K: F(3, 54) = 4.15, p = .01</td>
<td>MxF: F(6, 108) = 10.36, p &lt; .001</td>
</tr>
<tr>
<td>Neutral</td>
<td>M: F(1, 18) = 194.62, p &lt; .001</td>
<td>GxM: F(3, 54) = 5.70, p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>F: F(2, 36) = 3.89, p &lt; .05</td>
<td>ExK: F(6, 108) = 12.26, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>K: F(3, 54) = 5.70, p &lt; .01</td>
<td>ExMxF: F(6, 108) = 26.66, p &lt; .001</td>
</tr>
<tr>
<td>Overall</td>
<td>E: F(3, 54) = 13.82, p &lt; .001</td>
<td>GxE: F(3, 54) = 3.19, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>M: F(1, 18) = 188.68, p &lt; .001</td>
<td>GxF: F(2, 36) = 3.75, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ExM: F(3, 54) = 8.57, p = .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ExF: F(6, 108) = 2.23, p &lt; .05</td>
</tr>
<tr>
<td>(2) Pitch accent</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G: F(1, 18) = 5.97, p &lt; .05</td>
<td>ExF: F(6, 108) = 2.79, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>E: F(3, 54) = 14.56, p &lt; .001</td>
<td>ExKxM: F(6, 108) = 45.47, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>F: F(2, 36) = 5.42, p &lt; .01</td>
<td>ExK: F(6, 108) = 4.35, p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>K: F(1, 18) = 63.40, p &lt; .001</td>
<td>ExK: F(6, 108) = 9.25, p &lt; .001</td>
</tr>
<tr>
<td>(3) F&lt;sub&gt;0&lt;/sub&gt; variation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>G: F(1, 18) = 5.33, p &lt; .05</td>
<td>ExM: F(3, 54) = 3.45, p &lt; .05</td>
</tr>
<tr>
<td></td>
<td>E: F(3, 54) = 17.14, p &lt; .001</td>
<td>MxF: F(2, 36) = 7.80, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>F: F(2, 36) = 5.00, p = .01</td>
<td></td>
</tr>
</tbody>
</table>

*Note. G, group; F, focus; M, modality; E, emotion; K, keyword.*
REFERENCES


Pell, M. D. 1997a. An acoustic characterization of speech prosody in right-hemisphere-damaged patients: Interactive effects of focus distribution, sentence modality, and emotional context. [unpublished doctoral dissertation]


