AN ACOUSTIC CHARACTERIZATION OF SPEECH PROSODY IN RIGHT-HEMISPHERE-DAMAGED PATIENTS: INTERACTIVE EFFECTS OF FOCUS DISTRIBUTION, SENTENCE MODALITY, AND EMOTIONAL CONTEXT

Marc D. Pell
School of Communication Sciences and Disorders
McGill University, Montreal
February, 1997

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements of the degree of Ph.D.

0-612-30358-6

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

15. Interaction in F_0 range of emotion, sentence modality, and focus for (a) short and (b) long stimuli.

16. Interaction of emotion and focus for the short and long stimuli.

17. Interaction of group and emotion for the short and long stimuli.

18. Interaction of (a) group, emotion, and focus for short stimuli, and (b) group, emotion, and sentence modality for long stimuli.

19. Interaction of emotion and sentence modality for short stimuli.

20. Interaction of group, emotion, sentence modality, and focus for long stimuli spoken in a (a) neutral, (b) sad, (c) happy, and (d) neutral context.

21. Interaction of emotion, sentence modality, and focus on the perception of emphatic stress on the Baseline task.

22. Interaction of group, emotion, and focus on the perception of emphatic stress on the D-Neutral task.

23. Interaction of (a) group, sentence modality, and focus and (b) group, emotion, and sentence modality on the perception of emphatic stress on the F-Neutral task.

24. Overall accuracy of the NC, RHD, and LHD subjects in the Focus Perception condition as a function of focus position and task level.

25. Percentage of total stimuli identified by each group as correct, as a "placement" error, or as lacking emphasis ("neutral").

26. Interaction of group and emotion on the perception of emotional prosody on the Baseline task.

27. Interaction of group, emotion, sentence modality, and focus on the perception (a) sad, (b) happy, and (c) angry prosody on the D-Neutral task.

28. Interaction of emotion, sentence modality, and focus on the perception of emotional prosody on the F-Neutral task.

29. Interaction of (a) group and emotion, and (b) emotion and task in the Emotion Perception condition overall.

30. Percentage of stimuli identified by each group as conveying sad, happy, angry, or "neutral" prosody.

ABSTRACT

A review of the literature on speech prosody suggests that the right hemisphere may be crucial in expressing and perceiving prosodic information, although hypotheses concerning the underlying nature of this specialization remain disparate (e.g., Behrens, 1988; Ross, 1981; Van Lancker & Sidtis, 1992). To illuminate the right hemisphere's role in prosodic processing, and to explore the interaction between linguistic and emotional suprasegmental cues in speech production and perception, two experiments were conducted. In Experiment 1, utterances conveying three prosodic distinctions (emphatic stress, sentence modality, emotional tone) were elicited from normal (NC) and right-hemisphere-damaged (RHD) adults and then subjected to acoustic analysis. Results indicated that the intonation patterns produced by RHD patients were relatively normal in overall shape, but significantly restricted in fundamental frequency (F_0) variation relative to those produced by normal subjects. The RHD speakers also supplied fewer duration and F_0 cues to emphatic stress, and demonstrated aberrant control of speech rate and mean F_0 in expressing discrete emotions relative to the NC speakers. In Experiment 2, six receptive tasks in which the F_0 or duration parameters of prosodic stimuli were systematically altered, were presented to NC, RHD, and left-hemisphere-damaged (LHD) adults for linguistic or emotional identification. Results obtained for this experiment revealed that both the RHD and LHD patients were impaired in the recognition of emotional prosody, but that only the LHD patients were disturbed in perceiving linguistic specifications via prosodic cues. The outcome of both experiments is discussed with respect to current theories of the lateralization of prosodic processing.

RÉSUMÉ

Une recension de la littérature consacrée à la prosodie linguistique incite à croire que l'hémisphère droit joue un rôle crucial dans l'expression et la perception de l'information prosodique, bien que les hypothèses relatives au caractère sous-jacent de cette spécialisation restent disparates (voir Behrens, 1988; Ross, 1981; Van Lancker & Sidtis, 1992). Pour élucider le rôle que l'hémisphère droit joue dans le traitement de l'information prosodique et analyser l'interaction des signaux suprasegmentaux linguistiques et émotifs dans la production et la perception de la parole, nous avons réalisé deux expériences. Dans l'expérience n° 1, nous avons demandé à des sujets adultes normaux (SN) et des sujets adultes cérébrolésés droits (CLD) différents énoncés comportant trois distinctions prosodiques (accent d'insistance, modalité de la phrase, ton) qui ont ensuite été soumis à une analyse acoustique. Les résultats révèlent que, par rapport à celles produites par les sujets normaux, les structures d'intonation produites par les patients CLD sont relativement normales quant à leur forme générale, mais considérablement limitées quant à la variation de la fréquence fondamentale (F_0). Par rapport aux sujets témoins normaux, les locuteurs CLD produisent également moins de signaux (durée et F_0) contribuant à l'accent d'insistance et présentent des aberrations du contrôle du débit et de la F_0 moyenne lorsqu'ils doivent exprimer des émotions discrètes. Dans l'expérience n° 2, nous avons soumis des sujets adultes normaux et des sujets cérébrolésés droits (CLD) et gauches (CLG) à six tâches de perception consistant à identifier, sur les plans linguistique ou émotif, des énoncés dont on avait systématiquement modifié les paramètres de fréquence fondamentale (F_0) ou de durée des stimuli prosodiques. Les résultats de ces expériences indiquent que les patients CLD et CLG présentent une atteinte sur le plan de la reconnaissance de la prosodie émotive, mais que seuls les patients CLG ont de la difficulté à percevoir les particularités linguistiques à partir des signaux prosodiques. Les résultats des deux expériences sont analysés à la lumière des théories actuelles de la latéralisation du traitement de l'information prosodique.

ACKNOWLEDGEMENTS

I gratefully acknowledge the kind support of: Carol Leonard and Matthew Sams, for providing their voices in preparing the stimuli; Anita Shuper, Marie Leconte, Kathy Palesch, and Jennifer Hutner, for helping with data collection; McGill Translation Services, for translating the abstract; Drs. Rachel Mayberry and John Ryalls, for their insight and guidance as members of my program and dissertation committees; and Dr. Shari Baum, *Supervisor Emeritus*, for her wisdom, generosity, and flexibility throughout this process. The patients participating in this study are cordially acknowledged for giving so freely of their time and for their commitment to research. Finally, thanks go to Foti for so patiently indulging my ever-changing moods while completing this project. A portion of this work appears in Michael Lynch (Ed.), *The Cognitive Science of Prosody: Interdisciplinary perspectives* (In press). This research was enabled through doctoral fellowships received from the Natural Sciences and Engineering Research Council of Canada, the Fonds de la Recherche en Santé du Québec, and the McGill University Faculty of Medicine.

INTRODUCTION

Humans are thought to possess a unique, perhaps innate, capacity to utilize abstract sign systems to relay their physiological needs and internal psychological processes to conspecifics. This apparent predisposition for humans to acquire language may constitute a phylogenetic adaptation that has accorded selective advantage to our species; the priority given to language learning in early infancy would seem to exemplify the importance of attaining communicative proficiency for humans. Despite an aptitude to acquire linguistic forms early in life, the ability to establish effective (i.e., native-like) communicative systems declines during ontogeny, possibly as the result of brain-maturational constraints (Bever, 1981; Lenneberg, 1967). Understanding the biological mechanisms that regulate (and perhaps limit) our ability to acquire and use language has evoked considerable curiosity, as this endeavour promises to illuminate one of the primary features of human nature.

Scientific approaches to the study of human language have focused predominantly on how *segmental* aspects of the speech code are supported in verbal behaviour. Thus, human languages have been analyzed with respect to their phonological, syntactic, and semantic components, and the biological foundation of these operations has been explored. However, the expression of communicative intent in natural discourse is not confined to segmental features. Rather, obligatory alterations in *suprasegmental* parameters of the voice (i.e., the *melodic* structure of speech, or speech *prosody*) constantly serve to enhance, elaborate, or even contradict the meaning of segmental information.

In fact, although prosody has at times been conceived of as an ancillary construct in the study of language structure, the import of speech prosody may be more central to human communicative abilities than that of segmental features. Prosody is the first component of speech that is recognized and produced purposively by the human infant (Fernald, 1989; Kaplan & Kaplan, 1971; Lenneberg, 1967). The acquisition of "higher-order" abilities such as phonology and syntax proceeds due to the *facilitative* effect of the child's understanding of prosodic representations (Fernald & Simon, 1984; Morgan, Meier & Newport, 1987). Phylogenetically, the similarity of prosodic structures in human speech to the vocal signals of nonhuman primates (Cosmides, 1983; Hauser & Fowler, 1992; Ohala, 1984) suggests that prosody may constitute a more primitive communicative system that has given rise to (more recent) segmental forms over the course of our behavioural evolution (Cosmides, 1983). The very basic nature of prosody, both ontogenetically and phylogenetically, invites inquiry into the nature of its functions and its disorders, if we are to fully appreciate our biological propensity for verbal communication.

But precisely how is prosody characterized? What are its functions and units, its physical and perceptual dimensions? Cross-linguistic investigations have established the universality of prosody to human languages, identifying universal tendencies in the functions encoded by prosodic specifications, and in the physical manifestation of certain prosodic structures (Bolinger, 1978). However, other aspects of prosody are more subject to the linguistic conventions of different language communities and are not, therefore, definable in terms of specific, universal acoustic forms. These issues and their supportive scientific literature are elaborated in the proceeding discourse.

Prosody

Speech prosody is characterized by vocal alterations in fundamental frequency (F_0), duration, and intensity that extend beyond the domain of a single segment (Fry, 1970; Lehiste, 1970). Perceptually, modulation of these acoustic cues represents changes in the pitch, length, and loudness of the speech stream (Fry, 1958). Although the same parameters contribute to both the production and perception of prosodic events and are therefore highly related, the relative importance of individual cues in the physical stimulus and its auditory processing do not establish a direct one-to-one correspondence (e.g., Fry, 1958). Moreover, the weight accorded to individual prosodic cues in both receptive and expressive modalities has been shown to differ somewhat both inter-personally (Behrens, 1988; Denes & Milton-Williams, 1962; Lieberman & Michaels, 1962) and cross-linguistically (Beckman & Pierrehumbert, 1986; Bolinger, 1978; Fry, 1970; Hadding-Koch & Studdert-Kennedy, 1964; Halle & Vihman, 1991; Majewski & Blasdell, 1968). These data suggest that the acoustic parameters underlying prosody may each operate independently at times to mark specific communicative events.

Despite the potentially independent contribution of multiple features, fluctuations in F_0 are generally considered of utmost importance in the detection and transmission of prosodic information (Bolinger, 1955; 1958; Denes, 1959; Fry, 1955; 1958; Lieberman, 1960; Morton & Jassem, 1965; O'Shaughnessy, 1979; Swerts & Geluykens, 1993). Alterations in the temporal flow of an utterance also provide essential cues to prosodic meaning (Cooper, Eady & Mueller, 1985; Eady & Cooper, 1986; Klatt, 1976; Scherer, 1986), although some view temporal changes as perhaps secondary to F_0 change in prosodic marking (Lyberg, 1979). Changes in the amplitude or perceived intensity of speech are now deemed of relatively minimal prosodic significance (Behrens, 1988; Bolinger, 1958; Brown & McGlone, 1974; Morton & Jassem, 1965; Ross, 1988; Streeter, 1978; Turk & Sawusch, 1996). Despite hierarchical tendencies in cue use, the manner in which F_0 and other prosodic features combine to convey discrete messages varies according to the speaker's intentions and occasionally results in a "trading-off" in the relevance assigned to specific prosodic parameters (Beach, 1991; Brown & McGlone, 1974; McRoberts, Studdert-Kennedy & Shankweiler, 1995). Accordingly, the production and perception of prosody should be viewed as a dynamic process involving a *complex* of potentially independent vocal cues, of which changes in fundamental frequency may constitute the dominant feature in many cases.

One factor that exerts a strong influence on the composition of prosodic structures is the "functional role" encoded by suprasegmental information. Researchers differ in the number of functional levels they attribute to prosody (e.g., Monrad-Krohn, 1947, delineated four distinct communicative purposes subserved by prosody). Nonetheless, two major roles are commonly accepted: a propositional or *linguistic* function, and an affective or *emotional* function. Linguistic prosody serves to enhance or elaborate the propositional content of an utterance (e.g., highlighting the semantic value of elements within a sentence); as such, linguistic prosody constitutes part of a speaker's conventionalized knowledge about his or her language system (Bolinger, 1978). In contrast, emotional prosody represents the vehicle by which internal fluctuations in a speaker's affective state are conveyed vocally. Emotional signals are encoded with the propositional message *in tandem*, but are more tied to the speaker's internal psychological environment and therefore lie outside the domain of linguistic competence (Frick, 1985; Fry, 1970; Murray & Arnott, 1993).

The vocal correlates of emotional states are generally present and identifiable throughout the propositional message; indeed, normal listeners have recognized emotional content in speech segments as short as 60 milliseconds (Pollack, Rubenstein & Horowitz, 1960). However, consensus has not been reached as to what dimensions of emotional contours are psychologically relevant in the transmission and reception of vocal affect. One factor that likely serves as a strong indicator of emotional content is the *activation* level of the speaker

(Davitz, 1964b; Huttar, 1968; Ladd, Silverman, Tolkmitt, Bergmann & Scherer, 1985; Murray & Arnott, 1993; Pakosz, 1983; Scherer, 1986; Siegman, Dembroski & Crump, 1992). This hypothesis would predict that alterations in vocal parameters convey information regarding the degree of physiological arousal being experienced by the speaker, from very little (e.g., sorrow, depression) to considerable (e.g., anger).

Other potentially relevant dimensions of emotional meaning encoded by vocal features include the *valence* of the emotion represented (positive-negative), and whether the emotion was initiated by the speaker (e.g., contempt) or evoked by changes in the environment (e.g., surprise), a dimension variably referred to as *control* or *strength*. (For a more comprehensive discussion of this topic, the reader is referred to Murray and Arnott, 1993). Despite our rudimentary understanding of the units specified by emotional prosody, the vocal sequelae of the "primary" emotional states (e.g., sorrow, anger, happiness) are believed to show universal tendencies in production and perception (Frick, 1985; Scherer, 1986). This being said, the specific phonetic form of emotional contours in different languages, and the extent to which the expression of particular emotional content is permitted in discrete linguistic communities, differ somewhat cross-linguistically due to sociolinguistic conventions or "vocal display rules" (Kramer, 1964; Scherer, 1986).

Like emotional vocal cues, linguistic-prosodic features span segmental units of various lengths, such as the word or sentence. However, unlike emotional features which convey a relatively uniform message throughout an utterance, the communicative distinctions specified by linguistic prosody are closely linked to linguistic units of a specific size (e.g., emphasis is marked on only the word to be highlighted within a discourse). Thus, the *domain* of linguistic representation over which prosodic cues operate yields further distinctions that affect prosody use. The most highly localized domain over which linguistic-prosodic features are considered communicatively relevant is the *phonemic level*. Here, the articulatory influence of contiguous phonemes gives rise to highly localized changes in prosodic parameters (e.g., an emphasized vowel tends to be longer and exhibit a lower F_0 peak following a voiced rather than a voiceless consonant (Haggard, Ambler & Callow, 1970; Klatt, 1976)). These fine alterations in the phonetic form of adjacent phonemes remain outside the purview of the present discussion.

At the *lexical* or *word level*, prosodic cues lend prominence to select syllables or words within an utterance. These cues may be specified in the lexicon, contributing distinctive information about otherwise identical segmental strings (e.g., phonemic stress or tone). Alternatively, prosodic highlighting of a specific word within an utterance may indicate that the speaker accords increased semantic value to the selected item (e.g., emphatic stress or *focus*). The marking of prosody for emphatic or contrastive purposes has been shown to be highly related to the "givenness" or "newness" of the selected material; information that is assumed by the speaker to be "given" or known by the listener (or derivable from preceding contextual information) tends to be "de-emphasized" prosodically by the speaker, whereas "new" information is more likely to receive focus via prosodic highlighting (Eefting, 1990; Ferreira, 1993; MacWhinney & Bates, 1978; Nooteboom & Kruyt, 1987; O'Shaughnessy, 1979; Vande Kopple, 1982). Thus, in contrast to phonemic stress which is lexically-entrenched, emphatic stress reflects the conscious attempt of the speaker to lend saliency to items of relative importance within a *distribution* of intonational features. Accordingly, emphatic stress assumes intact *pragmatic* awareness on the part of the speaker of prior contextual cues and of the listener's knowledge.

Other suprasegmental parameters gain relevance at higher levels of linguistic representation. At the *phrase level*, prosodic discontinuities mark the presence of a syntactic break or juncture, an operation believed essential to efficient syntactic processing and the intelligibility of speech (Beckman, 1996; Bolinger, 1978; Lehiste, Olive & Streeter, 1976; Price, Ostendorf, Shattuck-Hufnagel & Fong, 1991; Wingfield, Lombardi & Sokol, 1984). At the *sentence* or *utterance level*, differences in the overall shape of the intonation contour signal the pragmatic intent of the utterance with respect to the listener (i.e., whether information is being requested, reported, etc.). This modal function of intonation contours may be linked to basic properties of speech production, such as the tendency of a speaker's fundamental frequency to decline over the course of an utterance (Cohen, Collier & 't Hart, 1982; Cohen & 't Hart, 1967; O'Shaughnessy, 1979; Pierrehumbert, 1979; Pike, 1945; cf. Lieberman, Katz, Jongman, Zimmerman & Miller, 1985; Umeda, 1982). This construct, alternately labelled "downdrift" or "declination", may constitute the "default" category by which the finality of utterances is signaled in human languages (Bolinger, 1978). Departures from this tendency (i.e., the absence of declination or a terminal rise in fundamental frequency) are often interpreted by the listener as an indication of something to follow, or a request to continue communicating (Bolinger, 1978; Ohala, 1984; Studdert-Kennedy & Hadding, 1973). Although some researchers stress the importance of only the terminal aspect of the intonation contour in distinguishing the modality of an utterance (Lieberman et al., 1985), it is more probable that further (albeit less relevant) modifications occur at other points of the intonation contour as well (Hadding-Koch & Studdert-Kennedy, 1964; Majewski & Blasdell, 1968; O'Shaughnessy, 1979; Studdert-Kennedy & Hadding, 1973).

As noted earlier, a fixed set of acoustic cues underlies linguistic and emotional prosody irrespective of the domain over which they are transmitted. Moreover, it is clear that the various propositional functions of prosody (e.g., contrastive highlighting, marking utterance modality) may be operative *simultaneously* in speech, as well as in conjunction with emotional signaling. As such, the need to relate information at one level of prosodic structure necessarily influences the parameters that cue information at other levels of representation (Fry, 1970; Lea, 1977; Ross, Edmondson & Seibert, 1986). With respect to the modulation of F_0 cues, for example, this "functional hierarchy" would appear to accord greatest weight to preserving local patterns of fluctuation at the phonemic and word levels (O'Shaughnessy & Allen, 1983). Locally-assigned features are then thought to be "superimposed" upon more global prosodic phenomena encompassing larger domains in speech planning, such as the phrase or sentence (O'Shaughnessy & Allen, 1983). The need for emotional expression exerts an additional influence on the prosodic parameters used to convey linguistic content (Fry, 1970; Gandour, Larsen, Dechongkit, Ponglorpisit & Khunadorn, 1995; Ross et al., 1986; Scherer, 1986). In fact, if sufficiently intense, the vocal correlates of emotional signals may impede any attempt to transmit propositional content simultaneously (Scherer, 1986).

Thus, the superposition of effects of sentence intonation, stress placement, and emotional content are of considerable theoretical interest if we are to understand how prosody is normally produced and perceived. Regrettably, theoretical accounts of how the various levels of prosodic structure interact in verbal behaviour have not been corroborated by appropriate acoustic descriptions. Although acoustic studies have explored the physical and perceptual correlates of *individual* prosodic functions, few have attempted to delineate the acoustic-perceptual underpinnings of prosodic messages when multiple constructs are present in the speech signal simultaneously. Indeed, preliminary work in this area constitutes one of the prime motivations of the present report. First, however, an overview of our current knowledge of the acoustic basis of discrete prosodic functions is presented, commencing with a discussion of emotional prosody.

Acoustic Investigations of Emotional Prosody

As stated above, F_0, duration, and intensity cues encode information about the affective state of the speaker or about the speaker's attitude towards the linguistic content of the utterance being spoken. For emotional communication, differences in *voice quality* (e.g., period-to-period fluctuations in F_0 or intensity) are believed to act as an additional defining feature (Bachorowski & Owren, 1995; Cummings & Clements, 1995; Ladd et al., 1985; Lieberman & Michaels, 1962; Murray & Arnott, 1993; Scherer, 1986). Unfortunately, little agreement currently exists as to how voice quality is defined or measured, limiting our understanding of the contribution of voice quality to affective speech.

In fact, despite numerous attempts to establish the acoustic underpinnings to discrete emotional meanings, the literature on vocal affect communication remains fragmentary, and acoustic descriptions of specific emotions remain elusive. This inability to identify the acoustic landmarks of basic emotions may arise from divergent views of what dimensions (and therefore measures) of emotional signals are communicatively relevant (see above). The lack of concordance in the acoustic literature on emotions may additionally reflect, at least in part, inter-personal differences in how emotional messages are normally transmitted and perceived (Ladd et al., 1985; Lieberman & Michaels, 1962; Pakosz, 1983).

What is generally agreed upon is that discrete emotions are not reflected in circumscribed emotional "contours" *per se*, but rather, that "global" modifications in the acoustic form of utterances contribute to distinctions in emotional content. The effects of global acoustic modifications occurring throughout the utterance may be particularly important in specifying emotions that differ as a function of the speaker's level of "activation". More precisely, acoustic investigations have illustrated that *increased* emotional tension is correlated with an increase in F_0 height and/or range, an increase in amplitude, and a faster rate of speech (Davitz, 1964a; Huttar, 1968; Scherer, 1974; Scherer, 1986). Increased pitch range over the course of an utterance is probably the most salient *auditory* determinant of a speaker's level of emotional involvement (Huttar, 1968; Ladd et al., 1985). The observation that emotions related in activation level (and not valence, for example) tend to be confused most frequently on perceptual tasks (Pakosz, 1983) demonstrates the importance of this psychological dimension for judging affect in speech, and suggests that emotions similar in activation correspond with systematic changes in the sound spectrum.

For example, the speaker's state of physiological arousal (or perceived involvement) may underlie some of the vocal differences between such emotions as "joy" and "anger" (both [+active]) and "sorrow" ([-active]): the former two emotions are expressed with a relatively high mean F_0 and high F_0 variability, whereas the latter is conveyed with a relatively low F_0 and little F_0 variation (Huttar, 1968; Pell & Baum, 1997b; Scherer, 1986; Van Lancker & Sidtis, 1992; Williams & Stevens, 1972). According to Huttar (1968) and others (Scherer, 1986), the tendency for F_0 and amplitude cues to be elevated in situations of increased emotional arousal stems from increased tension in the laryngeal and respiratory musculature. Certainly, it is recognized that increased muscular tension throughout the body is a concomitant of emotional activity (Ekman, Levenson & Friesen, 1983; Scherer, 1986).

Differences in emotional valence, or the perceived "pleasantness" or "unpleasantness" of prosodic attributes, also determine the acoustic manifestation of affective signals. In general, it has been shown that the adoption of a relatively high vocal pitch height, *in addition* to indicating involvement, contributes to the impression that a speaker is "pleasant" or "non-aggressive" (Davitz, 1964a; Huttar, 1968; Ohala, 1984; Uldall, 1960). In fact, the significance of this signal would appear to hold true for a large number of bird and mammalian species, suggesting that it may be biologically specified (Ohala, 1984). The increase in mean F_0 and F_0 variability described for "joy" or "happiness" may therefore encode distinct information about both the speaker's activity level and the valence of his or her emotional disposition (Fonagy, 1978; Huttar, 1968; Lieberman, 1961; Scherer, 1974). Emotions differing in one dimension but similar in another (e.g., joy and anger are similar in activation but not in valence) must be distinguished by other prosodic means; given the present example, it is likely that differences in the *rate* of F_0 change play a role, with anger demonstrating rapid F_0 excursions on stressed syllables and joy demonstrating gradual, "smoother" transitions across syllables (Fairbanks & Pronovost, 1939; Fonagy, 1978; Williams & Stevens, 1972). Alterations in voice quality have also been offered as highly pertinent to the attribution of positive and negative emotional states (Johnson, Emde, Scherer & Klinnert, 1986; Ladd et al., 1985; Scherer, 1986), although, as noted above, these parameters remain poorly specified.

From the foregoing discussion, it seems likely that vocal cues representing several independent psychological dimensions interact in a rather complex fashion to transmit discrete emotional messages. Although it is clear that changes in each of the three primary acoustic cues (F_0, duration, amplitude) normally facilitate the communication of emotional distinctions (for a hypothetical account of how each variable is modulated to convey various emotions, see Scherer, 1986), it is similarly obvious that fundamental frequency change throughout the utterance constitutes the dominant cue (Bolinger, 1978; Lieberman & Michaels, 1962). Further inquiry may help reconcile some of the uncertainties currently characterizing this body of literature. First, however, our discussion of the acoustic basis of prosody examines how *linguistic* functions are represented in speech production and perception.

Acoustic Investigations of Linguistic Prosody

A greater number of investigations have explored the acoustic basis of linguistic prosody in normal speakers (e.g., Cooper et al., 1985; Eady & Cooper, 1986; Farnetani, Taylor Torsello & Cosi, 1988; Lieberman, 1960; McRoberts, Studdert-Kennedy & Shankweiler, 1995; O'Shaughnessy & Allen, 1983; Ryalls, Le Dorze, Lever, Ouellet & Larfeuil, 1994; Weismer & Ingrisano, 1979).

Accordingly, much is now understood about the manner in which word and sentence level prosodic phenomena convey linguistic distinctions. Syllable or word stress, as outlined above, assumes a dual role in propositional communication: one role defined by linguistic conventions (phonemic stress), and another more pragmatic role reflecting the speaker's on-line attempt to enhance the semantic value of elements within an utterance (emphatic stress). Acoustically, these two communicative functions are manifested in the speech signal in a highly similar, albeit distinct manner.

Specifically, numerous studies have illustrated that the F_0 of a word is significantly higher when "stressed" than "unstressed" (Brown, Strong & Rencher, 1974; Cooper, Soares, Ham & Damon, 1983; Eady & Cooper, 1986; Fry, 1958; Lieberman, 1960; McRoberts et al., 1995; Morton & Jassem, 1965; Ohala, 1977; O'Shaughnessy, 1979). The tendency to elevate the F_0 of stressed syllables in most languages may be important in facilitating the more basic feature of word stress, F_0 *excursions* around the marked syllable (Bolinger, 1958; Cohen & 't Hart, 1967; Lea, 1977; Morton & Jassem, 1965; O'Shaughnessy, 1979). These "pitch accents" may be encoded as either rises or sharp falls in F_0, indicating that the *extent* of F_0 change, and not the *direction* of change, is central to cueing the stress feature (Bolinger, 1958; O'Shaughnessy, 1979). The *relative magnitude* of local F_0 excursions would appear to be the parameter that determines whether stress cues are being used emphatically in speech: emphatic stress tends to be associated with much larger F_0 obtrusions than those encoded by phonemic stress (Cooper et al., 1985; Eady & Cooper, 1986; O'Shaughnessy, 1979).

In addition to pitch accents, temporal modifications are also utilized to highlight linguistic prominence on select syllables or words. In general, stressed syllables tend to be longer in duration than unstressed syllables (Klatt, 1976). Furthermore, emphasized words are generally longer than the same word without emphasis (Eady & Cooper, 1986; Eefting, 1990; Ferreira, 1993; Fry, 1955; Klatt, 1976, McClean, 1973 #227; Morton & Jassem, 1965; Weismer & Ingrisano, 1979), although not invariably (Cohen & 't Hart, 1967). Lyberg (1979) has argued that increased vowel durations on focused elements represent a secondary effect to accomodate (more relevant) alterations in F_0 events. However, contextual factors such as sentence position have been shown to bring about a "trading relationship" between F_0 and duration, for which temporal parameters become the *primary* cue to emphasis at different points in the utterance (Brown & McGlone, 1974). Thus, the durational features of linguistic focus can not be described as wholly redundant with F_0 cues under all conditions in normal speech.

The acoustic underpinnings of *intonational* distinctions, as described earlier, rely strongly on the direction of the terminal portion of the F_0 contour, a terminal fall signifying finality (declaration) and a terminal rise indicating lack of finality (or interrogation). Although a terminal rise in F_0 is the most common device used to mark interrogation in human languages (Bolinger, 1978; Hermann, 1942, cited in Bolinger, 1978), this pattern is not invariably present in utterances recognized as interrogatives. Rather, a comparatively high F_0 throughout the utterance or at strategic points in the utterance may be sufficient to indicate that a question is intended (Hadding-Koch & Studdert-Kennedy, 1964; Majewski & Blasdell, 1968; O'Shaughnessy, 1979; Studdert-Kennedy & Hadding, 1973).

In fact, several researchers have stressed the importance of a "turning point" in the F_0 contour which may be crucial in distinguishing declarative and interrogative utterances; their evidence suggests that when the turning point (which may differ from subject to subject) is perceived by the listener as relatively high in pitch, the effect of the rising or falling terminal is "overriden" and the utterance is recognized as a question (Hadding-Koch & Studdert-Kennedy, 1964; Majewski & Blasdell, 1968; Studdert-Kennedy & Hadding, 1973). Thus, the acoustic-auditory determinants of speech mode distinctions are present in the entire F_0 contour, despite the paramount importance of the terminal. Ethologists have surmised that the use of a rising F_0 pattern to signal questions is consistent with the perceived "pleasantness" and "submissiveness" of rising contours (see above); by requesting information, the speaker seeks the cooperation and goodwill of the listener, an act best facilitated through the adoption of a pleasant, submissive speaking tone (Ohala, 1984). Although only speculative, this suggestion underlines the potential insights to be gained from a better understanding of how linguistic and emotional representations interact in vocal behaviour.

Relatively few studies have explored how acoustic modifications in the speech stream accomodate linguistic meanings at multiple levels of prosodic structure concurrently. Preliminary work in this area was accomplished in a series of inter-related studies conducted by Cooper, Eady, and their colleagues (Cooper et al., 1985; Eady & Cooper, 1986; Eady, Cooper, Klouda, Mueller & Lotts, 1986). In an initial investigation, the authors utilized an elicitation paradigm to explore how changes in two parameters--F_0 and duration--were influenced by the position of emphatic stress within declarative utterances (Cooper et al., 1985). In a subsequent study, the same procedure was employed to assess changes in the F_0 and duration parameters underlying emphasis in both declarative *and* interrogative sentences (Eady & Cooper, 1986). Through these data, the authors were attempting to establish a preliminary understanding of how word- and utterance-level (linguistic) prosodic features interact in the speech signal.

In their initial investigation, Cooper et al. (1985) required six normal male subjects to produce declarative sentences in which emphatic stress was placed at one of four distinct "keyword" positions within the utterance: sentence-initial, sentence-medial (two positions), and sentence-final (e.g., The *ship* is *departing* from *France* on *Sunday*). Measures of duration and peak F_0 were determined for each keyword of each utterance (keywords were those content words that could plausibly receive focus). Subsequently, the four versions of each utterance were compared with respect to the acoustic features. In this manner, the authors hoped to define the acoustic correlates of linguistic focus as a function of sentence position, and additionally, determine the effects of localized prosodic phenomena at remote points in the utterance (i.e., their effect on the duration values and F_0 "topline" of the intonation contour).

The results demonstrated that for duration, emphatic stress was always expressed by means of a significant increase on the focused item, when compared to the same item in the same position without focus (Fry, 1955; Klatt, 1976, McClean & Tiffany, 1973; Morton & Jassem, 1965; Weismer & Ingrisano, 1979). This increase in the duration of focused elements was significant in all four sentence locations, albeit much smaller in sentence-final position (16% increase sentence-finally vs. 40% increase in first three positions). The duration of *unfocused* words was unaltered by the location of focused information in the utterance. This latter finding was interpreted by the authors as evidence that the temporal correlates to sentence focus are likely confined strictly to the item that receives emphasis in speech production (Cooper et al., 1985; cf. Weismer & Ingrisano, 1979).

With respect to concomitant changes in fundamental frequency, the data revealed that emphasis was not marked by higher F_0 on focused words, but rather, by a sharp decline in F_0 on *subsequent* items. In sentence-final position where a post-focus drop in F_0 was not possible, the final word was shown to be (non-significantly) higher in the focused condition than in the other three renditions of the utterance when the item was unfocused. Unlike duration, therefore, the effects of emphatic stress on F_0 were not strictly localized to the emphasized item, but were evident on post-focus items as well. Indeed, normal speakers appeared to "deemphasize" all material that occurred subsequent to the focused word by means of substantially lowered F_0 values. The fact that large F_0 excursions were not evident on focused words in sentence-final position, and that lengthening due to focus is diminished in this position as well, led the authors to conclude that speakers do not typically mark focus in sentence-final position as distinctively as in other locations within the utterance (Cooper et al., 1985).

To elaborate upon these findings, Eady and Cooper (1986) extended their paradigm to include both declarative and interrogative sentences, employing a subset of the stimuli analyzed by Cooper et al. (1985). Six additional normal speakers were asked to produce the test stimuli (n=4) as both a statement and a question, and by placing emphatic stress at one of three separate locations within the utterance for each sentence type. The authors further elicited "neutral" versions of each utterance (i.e., sentences without a focused word) to better characterize the acoustic attributes of emphatic stress relative to sentences without linguistic focus. As conducted previously, duration and peak F_0 measures were calculated at each keyword position for each utterance.

The results confirmed those of Cooper et al. (1985) with respect to the acoustic cues to emphatic stress in declarative sentences: focus was marked through increased duration of the emphasized word followed by a rapid post-focus decrement in F_0. Again, these cues were diminished in sentence-final position where the increase in syllable lengthening was reduced and a post-focus F_0 fall could not occur. In fact, for fundamental frequency, utterances with sentence-final focus did not significantly differ from neutral sentences without focus (Eady & Cooper, 1986). This outcome is in accordance with previous data indicating that the acoustic attributes of emphasis are context-dependent, varying as a function of the serial position of the focused item within the utterance (Brown & McGlone, 1974).

Finally, in comparing corresponding declarative and interrogative utterances, no differences were noted with respect to the duration of either focused or unfocused words. The only exception was observed for sentence-final words in neutral questions (without focus) which tended to be 11% longer than the same words in neutral statements. For fundamental frequency, however, sentence-initial focus in questions resulted in elevated F_0 peaks on all subsequent words in the utterance in contrast to the greatly lowered F_0 values for the same words in statements with sentence-initial focus (Eady & Cooper, 1986). These data reiterate the importance of F_0 in marking the declarative/interrogative distinction in spoken English (Bolinger, 1978; O'Shaughnessy, 1979). More importantly, they reveal that the cues underlying the sentence modality distinction are markedly (and systematically) influenced by other suprasegmental processes such as sentence focus (Eady & Cooper, 1986).

Based on their collective findings (Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986), it is suggested that both highly localized (for duration) and more global (for F_0) prosodic features contribute simultaneously and independently in communicating emphasis in speech. These data are generally corroborated by the literature on linguistic focus when F_0 cues are considered (see Behrens, 1988; McRoberts et al., 1995; O'Shaughnessy, 1979; Ryalls et al., 1994). As noted previously, Cooper and Eady's data are not always consistent with descriptions of the temporal aspects of focus, as duration changes have been noted in *nonemphasized* segments by other researchers (Weismer & Ingrisano, 1979). However, as argued by Eady & Cooper (1986), utterance length may constitute an additional factor affecting the extent to which acoustic modifications are utilized to signal emphasis in speech; they argued that only long utterances (i.e., more than 8 syllables, ressembling those presented by Eady & Cooper, 1986) may demonstrate temporal changes isolated to the focused element, whereas short utterances (i.e., approximately 5 syllables, ressembling those presented by Weismer & Ingrisano, 1979) show less localized effects. Thus, the potential influence of utterance length on the cues to focus constitutes an additional factor meriting more in-depth study.

Eady & Cooper's (1986) observation that the cues to emphatic stress influence those that convey the sentence modality of utterances (declarative/interrogative) is in accord with recently published data (McRoberts et al., 1995). In that study, the ability to modulate F_0 to signal both emphatic stress and interrogation on the final syllable of short utterances (e.g., *November*) led to a trading relationship, whereby an increase in prominence due to emphasis resulted in a decrease in F_0 rise due to interrogation. The same utterances spoken in a "positive" or "negative" affect did not affect the size of the terminal F_0 glide, suggesting to the authors that the control of F_0 for linguistic-prosodic distinctions may be functionally separate from that of affective distinctions (McRoberts et al., 1995).

Thus, preliminary acoustic evidence that a "trade-off" of some sort may occur when the cues to word- and sentence-level linguistic contrasts are used simultaneously has begun to surface (Eady & Cooper, 1986; McRoberts et al., 1995). Overall, this line of research is valuable in elucidating how multiple prosodic cues may be simultaneously realized in speech production. In a related manner, it suggests what auditory dimensions may be important in the *perception* of prosodic content in natural discourse. Unfortunately, the data reported by Cooper, Eady and their colleagues are inadequate as a model of how suprasegmentals are utilized in natural speech production, as they fail to consider how vocal affect is integrated with propositional content in the speech signal. McRoberts et al.'s (1995) findings begin to address this question, but remain highly speculative at this time and await corroboration using a larger range of test stimuli and acoustic measures.

Finally, none of these investigations sought to address how the production and perception of prosodic messages are subserved or lateralized in the brain. Much may be learned by studying *disorders* of speech prosody and their potential association with aspects of prosodic structure. This topic--the neural mechanisms underlying our capacity for emotional and linguistic prosody--is explored in detail in the following chapter.

NEURAL SUBSTRATES OF PROSODY

The observation that acquired damage to the central nervous system may culminate in a specific disorder affecting the "prosodic quality of speech" is credited largely to the Norwegian neurologist Monrad-Krohn. He described a patient who was rendered 'dysprosodic' following a shrapnel wound to the left fronto-temporo-parietal region of her brain; as a result of this injury, alterations in the woman's speech led her to be perceived as having a foreign (possibly German) accent, despite her Norwegian heritage (Monrad-Krohn, 1947). Since that time, considerable interest in the effects of acquired disease of the brain on the ability to produce and comprehend speech prosody has been generated (see Baum & Pell, In press, for a review). Through the association of specific regions of neurological dysfunction and the failure to modulate aspects of prosody, this line of inquiry has sought to illuminate the neural mechanisms underlying our "prosodic faculty" with greater precision. Ultimately, refining our knowledge of how the brain regulates *suprasegmental* aspects of communication may inform current theories on the neurological basis of linguistic communicative abilities, and lead to a more complete account of the human propensity for verbal behaviour.

However, investigations that attempt to correlate behavioural deficits with circumscribed areas of brain damage must be approached with some caution. For example, deficits seemingly ascribable to focal neuroanatomical lesions may in some cases reflect functional damage to brain structures or pathways only *indirectly* related to the focal neuroanatomical site. This possibility may yield incorrect assumptions about the role of discrete brain regions in the modulation of behavioural functions, such as prosody. When viewed across studies, variability in the clinical attributes of patients (e.g., differences in lesion site or size, acuteness of injury) and in the assessment procedure utilized are frequently present; these methodological irregularities may render the comparison of otherwise complementary studies problematic, if not inappropriate. Collectively, these issues present a challenge to those who wish to gain insight into the neural representation of prosodic functions via lesion analysis.

Despite these caveats to interpreting the neurolinguistic literature on prosody, controlled examination of the relative sparing and loss of prosodic functions in brain-damaged individuals constitutes a viable means of exploring brain-prosody relationships. Indeed, this methodology has permitted considerable advancement in our knowledge in this area over the past couple of decades. This being said, much remains uncertain or ill-specified about how prosodic phenomena are represented in the brain. A review of the relevant literature begins with research exploring the *production* of prosody in neurologicallyimpaired subjects, followed by reports addressing *receptive*-prosodic capabilities in brain-damaged individuals.

Production of Prosody

It has been recognized for some time that disorders affecting the motor control of the speech musculature routinely lead to articulatory defects, including various alterations in the prosodic quality of speech (Darley, Aronson & Brown, 1969; Kent & Rosenbek, 1982). However, the association of focal brain-damage and expressive disturbances of prosody in the *absence* of motor speech impediments was borne largely by clinical impressions of flattened or restricted prosodic variation in the speech of *right-hemisphere-damaged* (RHD) patients. Although clinical reports initially focused exclusively on the failure of RHD patients to impart appropriate affective features to their speech, more recent research suggests that expressive-prosodic abnormalities may not always be tied to an affective context, and may become manifest in both RHD *and* LHD populations. These issues are considered in the ensuing discussion, commencing with investigations of emotional prosody production in brain-injured adults.

Emotional Prosody

As cited above, a number of investigators have explored the effects of unilateral right hemisphere dysfunction on the ability to intone speech for *emotional* purposes. In an early study, Tucker, Watson, and Heilman (1977) required 8 RHD patients with parietal disease and neglect and 8 neurologically-intact control subjects to repeat emotionally-neutral sentences (e.g., *The boy went to the store*) in various affective tones (happy, sad, angry, indifferent). Recordings of the utterances were then presented in random order to three raters, who judged the accuracy of each subject in conveying the target emotions. Results of the perceptual ratings demonstrated that non-brain-damaged control subjects were significantly better able to signal emotional meanings through prosody than RHD patients, who performed at chance level overall. The authors concluded from these findings that patients with right temporoparietal lesions and neglect exhibit a defect in the production of affective prosody (Tucker, Watson & Heilman, 1977). However, the authors cautioned that the evocative task they employed may not have fully characterized the ability of RHD patients to impart emotions in natural, *spontaneous* speech.

In a number of case reports, Ross and his colleagues (Gorelick & Ross, 1987; Ross, 1981; Ross, Harney, deLacoste-Utamsing & Purdy, 1981; Ross & Mesulam, 1979) have pursued the hypothesis that right hemisphere dysfunction is critical in producing expressive disturbances of emotional prosody. Initially, Ross and Mesulam (1979) supplied anecdotal evidence of two patients with right-hemisphere lesions and "flat affect" (i.e., restricted prosodic variation), but conducted no formal testing to substantiate their impressions. In two subsequent studies (Ross, 1981; Ross et al., 1981), additional cases of RHD patients with putative difficulties in producing or comprehending emotional prosody were described. In these latter investigations, perceptual impressions of each RHD patient's speech were garnered by one of the authors following a brief bedside evaluation, establishing the presence or absence of a prosodic defect (or *aprosodia*, as coined by Ross, 1981).

Despite the relatively uncontrolled manner in which their data were collected, the authors viewed the cases reported in this series of studies as powerful evidence of the right hemisphere's superiority in the modulation of affective language, including emotional prosody and gesture (Ross, 1981; Ross et al., 1981; Ross & Mesulam, 1979). Ross (1981) further postulated--again, based on this rather small number of case reports-- that the affective components of language dissociate in a manner analogous to the aphasias (e.g., *motor aprosodia*, *conduction aprosodia*). Furthermore, he posited that the neural organization of the aprosodias in cortical regions of the right hemisphere mirrors that of the aphasias in the left hemisphere (e.g., an anterior, fronto-parietal lesion would lead to *motor aprosodia*, or specific expressive deficits for affective prosody, akin to Broca's aphasia). The reported failure of Ross and his colleagues to encounter a "negative" case of a RHD patient whose aprosodia did not conform to the lesion site predicted by Ross (1981) was cited as validation of this hypothesis. However, it is noteworthy that two RHD individuals without the anticipated deficits were classified as "crossed aprosodics" due to their incompatibility with this scheme (Gorelick & Ross, 1987).

In fact, there is sufficient cause to view many of the conclusions arrived at by Ross and his associates as tentative, at least in their strong form (see Ryalls, 1988). As intimated above, the failure of these researchers to supply sufficiently controlled measures of their patients' prosodic deficits (data were based on the subjective intuitions of one of the authors) poses a general concern about the reliability of their diagnoses. Indeed, perceptual judgements of prosody have been shown susceptible to bias even in trained speech professionals (Alexander & Bakchine, 1994; Lieberman, 1965). With regard to Ross' (1981) "mirror hypothesis", numerous production studies have now reported patients with prosodic disturbances that clearly do not conform to this hypothetical model (Borod, Koff, Perlman Lorch & Nicholas, 1985; Brådvik et al., 1990; Brådvik et al., 1991; Cancelliere & Kertesz, 1990; Lebrun, Lessinnes, De Vresse & Leleux, 1985; Tucker et al., 1977; Weintraub, Mesulam & Kramer, 1981). Finally, the more general claim that right hemisphere structures are dominant for emotional prosody production, although plausible, was founded on analyses that ignored the potential contribution of *left-hemisphere* mechanisms on such tasks (Ross, 1981; Ross et al., 1981; Ross & Mesulam, 1979; also Tucker et al., 1977). Certainly, the lateralization of a behavioural function cannot be ascertained until the role of *both* cerebral hemispheres is clearly defined, an obvious shortcoming of these investigations.

In an investigation of emotional prosody production in both RHD *and* LHD patients (and healthy control subjects), Borod, Koff, Lorch, and Nicholas

(1985) elicited samples of emotional speech from subjects by requiring them to comment on pleasant or unpleasant images (e.g., a beautiful sunset or a victim of starvation). Two raters then graded the extent to which each group utilized intonation and other nonverbal cues in their responses, based on a four-point scale. The results indicated that the RHD patients were perceived to use emotional prosody significantly "less often" than the LHD patients in these emotionally-laden situations; curiously, however, neither clinical group differed significantly from the normal group on this measure. Weak evidence that the right hemisphere is relatively more involved than the left hemisphere in the ability to produce emotional prosody may therefore be inferred from these data (Borod et al., 1985; Ross et al., 1981). It is additionally noteworthy that the performance of the RHD patients was not shown to differ as a function of lesion site (prerolandic vs. postrolandic), inconsistent with Ross' (1981) assertion that only *anterior* (i.e., prerolandic) right hemisphere lesions disturb the production of affective prosody.

More recent studies have benefited from acoustic analyses of patients' speech--a less subjective measure than perceptual ratings--to explore the nature of prosodic abnormalities following brain injury. For instance, Edmondson, Chan, Seibert, and Ross (1987) required 8 RHD Taiwanese speakers with fronto-parietal lesions and 8 healthy control subjects to repeat sentences spoken in various emotional tones. Each group's productions were then analyzed both perceptually and acoustically. Perceptual ratings indicated that the RHD patients were significantly less accurate than the control subjects in utilizing prosodic features to signal discrete emotions (the RHD patients' speech was described as emotionally "flat" relative to that of the control subjects). Interestingly, significant between-group differences in the modulation of several acoustic measures, in particular, an attenuation of F_0 variation in the RHD subjects, appeared to corroborate the perceptual data. These findings were interpreted as cross-linguistic evidence that emotional prosody may be modulated dominantly by the right hemisphere (Edmondson, Chan, Seibert & Ross, 1987). More ambitiously, the authors conjectured that their evidence of acoustic abnormalities in RHD Taiwanese speakers served to support previously-documented, anecdotal reports of RHD English-speaking aprosodics (Ross, 1981; Ross et al., 1981; Ross & Mesulam, 1979; but cf. Borod et al., 1985; Cancelliere & Kertesz, 1990; Tucker et al., 1977).

In a related but more recent study, Gandour and his colleagues (Gandour et al., 1995) conducted perceptual and acoustic analyses of emotional speech elicited from RHD Thai speakers. Based on perceptual indices, the ability of the RHD patients to convey affective qualities in their speech was again shown to be deficient relative to that of non-neurological control subjects (Edmondson et al., 1987). However, contrary to Edmondson et al's (1987) findings, few substantial differences emerged between the RHD and control subjects in the ability to modulate the acoustic underpinnings to emotional messages, including no between-group differences in how F_0 cues were manifested (Gandour et al., 1995). The observation that both groups patterned similarly in producing the acoustic correlates to discrete emotions suggested that impairments of emotional prosody production subsequent to right-brain-damage may be *quantitative*, rather than qualitative, in nature (Gandour et al., 1995). Coupled with previous data (Behrens, 1988; Gandour et al., 1992), the authors intimated that the *domain* over which prosodic cues operate, and not simply their linguistic or affective function, may be a critical determinant of prosody lateralization, with larger prosodic units (such as sentence intonation) showing a stronger right hemisphere bias than smaller units (e.g., phonemic stress or tone).

To further illustrate the RH's role in the production of emotional prosody, Ross, Edmondson, Seibert, and Homan (1988) examined the ability of 5 right-handed *normal* subjects to convey affective prosody before, during, and after a Wada (sodium amytal) test, administered to the subjects' right hemisphere. Recordings of each subject's speech were made at each stage of the experiment and then subjected to acoustic analysis. Results demonstrated that *during* the Wada procedure (i.e., while the right hemisphere was temporarily "deactivated"), subjects were unable to appropriately intone their voices for affective purposes. Although subjects displayed affectively "flat" speech during the test, no acoustic differences were noted pre- or post-Wada (Ross, Edmondson, Seibert & Homan, 1988). These findings speak persuasively to an important role for the right hemisphere in the production of emotional prosody. However, contrary to the authors' contentions, these data fail to reconcile the *exclusivity* of right hemisphere participation with emotional prosody, as similar analyses were not undertaken during a *left*-sided Wada test.

In a much cited study (Shapiro & Danly, 1985), RHD, LHD, and non-brain-damaged control subjects read target sentences (e.g., *He will be here tomorrow*) within contexts designed to bias both emotional and *linguistic*-prosodic interpretations at the sentence level (happy *vs.* sad; declarative *vs.* interrogative, respectively). Acoustic measures of each target utterance were then extracted to characterize the ability of each clinical group to inflect the voice for specific affective and linguistic purposes. Results indicated that the RHD patients exhibited significantly less pitch variation and intonational range in both affective *and* linguistic contexts relative to the LHD and healthy subjects. Moreover, the "type" of dysprosody observed in RHD patients was shown to vary as a function of lesion site; specifically, anterior and central RHD resulted in *restricted* intonational range, whereas posterior (postrolandic) RHD led to *exaggerated* pitch variation and range (Shapiro & Danly, 1985). Based on these findings, the authors concluded that right hemisphere insult alone results in a "primary disturbance" of speech prosody that is not tied to its affective components.

However, the validity of many of Shapiro and Danly's (1985) conclusions is suspect on several fronts. Their proposal that anterior and posterior right hemisphere lesions yield "opposing" types of dysprosody was based on a very limited subject sample (6 anterior/central and 5 posterior RHD patients) and the authors themselves remarked on considerable individual variability within each of these (already small) groups. Furthermore, the failure of Shapiro and Danly to normalize their acoustic data (particularly F_0 range and variability measures) for differences in speaker mean F_0 prior to analysis may have led to an incorrect attribution of "hypermelodicity" in some of their RHD patients (Colsher, Cooper & Graff-Radford, 1987; Ryalls, 1986). Finally, as pointed out by Ryalls (1986), Shapiro and Danly's (1985) assertion that RHD *alone* leads to an "intrinsic deficit in the modulation of speech prosody" is incompatible with the same authors' previous work demonstrating qualitatively similar impairments in the production of prosody by LHD aphasic patients (Danly, Cooper & Shapiro, 1983; Danly & Shapiro, 1982). These discrepancies and methodological shortcomings diminish the strength of many of Shapiro and Danly's (1985) conclusions. Nonetheless, their proposal that prosodic impairments following RHD may extend to both affective and non-affective stimuli remains an intriguing prospect for further study.

To more precisely explore the relationship between intrahemispheric lesion location and disturbances of emotional prosody production and comprehension, Cancelliere and Kertesz (1990) assessed 46 patients with unilateral infarcts (28 RHD, 18 LHD) and 20 neurologically-intact control subjects on a standardized battery of emotional prosody tests. In their study of production abilities, one task required subjects to intone "neutral" sentences in various affective tones and a second task required subjects to simply repeat utterances that were already emotionally-charged (three raters judged the accuracy of the emotions conveyed).

The performance of each patient was subsequently classified according to the aprosodic syndromes outlined by Ross (1981) and the CT scans for patients with corresponding aprosodias were superimposed in an attempt to correlate their aprosodia with lesion site information.

Perhaps surprisingly, the results of the CT overlap technique (based on both production and comprehension subtests) revealed that a proportionate number of RHD (75%) and LHD (78%) patients were classified as aprosodic overall (4 RHD and 2 LHD patients manifested a specific "motor" aprosodia). These data, therefore, fail to substantiate claims that only RHD disturbs the production and comprehension of emotional speech (e.g., Ross, 1981; Tucker et al., 1977). Further analyses undertaken by the authors involved reclassifying the patients into broadly-defined groups according to lesion site (e.g., right or left anterior, central, posterior, etc.) and comparing their performance on specific prosody subtests. Interestingly, these analyses again revealed no significant differences for *any* subtest as a function of either intra- or inter-hemispheric lesion site. These data, derived from more stringent, objective measures than those of previous reports (Ross, 1981; Ross et al., 1981; Ross & Mesulam, 1979), argue compellingly that both hemispheres may be engaged in the production and comprehension of emotional aspects of speech (Cancelliere & Kertesz, 1990). The involvement of basal ganglian dysfunction in producing aprosodic syndromes was specifically highlighted by their data, suggesting that subcortical structures may be of considerable import in the neural control of emotional prosody.

Not all research has demonstrated an association between focal right-hemisphere lesions and defects in the production of emotional prosody. Hird and Kirsner (1993) found no significant acoustic differences between RHD and healthy control speakers in the ability to produce four distinct emotions (happy, sad, angry, neutral), but their measures were limited to *temporal* aspects of the subjects' speech. Relatedly, a RHD patient described by Lebrun et al. (Lebrun et al., 1985) demonstrated intact expression of emotional prosody following a right temporoparietal excision. However, standardized tests were not used in this assessment.

Finally, two inter-related studies conducted by Brådvik and his colleagues (1990; 1991) failed to uncover deficits in the production of emotional prosody in relatively sizable groups of RHD Swedish-speaking individuals. Curiously, Brådvik et al. (1991) did observe difficulties in their RHD patients' ability to mark certain *linguistic*-prosodic distinctions, such as the ability to signal differences in "speech acts" (e.g., statements, questions, commands). Although it is not immediately clear why their RHD patients exhibited expressive impairments for linguistic but not emotional intonation (the RHD patients' *comprehension* of both emotional and linguistic prosody was impaired), Brådvik et al.'s data again question whether right hemisphere participation in prosody encoding is confined to affective cues (Shapiro & Danly, 1985).

Most recently, Baum & Pell (1997) employed acoustic analysis to study the production of emotional and linguistic prosody by RHD and non-brain-damaged adults, as well as a small control group of LHD patients. Both repetition and reading tasks were utilized to elicit sentences differing in emotional and linguistic content. For both repetition and reading tasks, subjects were further required to model stimuli varying in the amount of linguistic structure provided (stimuli were semantically well-formed and emotionally-biased, composed of nonsense syllables, or filtered of the phonetic content).

In general, results indicated that both the RHD and LHD patients were able to manipulate the acoustic underpinnings of emotional- and linguistic-prosodic messages in a manner similar to that of normal speakers (Baum & Pell, 1997; see also Gandour et al., 1995; Shapiro & Danly, 1985). This finding suggests that RHD patients may have largely retained the ability to impart affective (and linguistic) features to their speech via prosody, contrary to previous contentions (e.g., Ross, 1981; Tucker et al., 1977). However, the authors did note some irregularities in the global control of F_0 in their clinical subjects (e.g., the RHD group exhibited a somewhat restricted F_0 range relative to the control group in some conditions). These abnormalities may indicate a right hemisphere bias for the control of F_0 in prosodic signaling, irrespective of the *functional* attributes of F_0 cues in speech (Baum & Pell, 1997).

As illustrated in the foregoing discussion, despite significant attention to investigating the capacity of the right cerebral hemisphere in the production of emotional prosody, little consensus has been reached in this literature. Although considerable data advocate a substantial role for the right hemisphere in this function, the failure of investigators to adequately address the potential of left hemisphere mechanisms for emotional prosody production impedes strong conclusions about the *dominance* of the right hemisphere for this function. Moreover, reports indicating that RHD patients may be disturbed in the production of non-affective prosody as well (Baum & Pell, 1997; Brådvik et al., 1991; Shapiro & Danly, 1985) exemplify the need for further definition of the RH's role in prosodic encoding. To this end, a thorough examination of how *linguistic*-prosodic contrasts are lateralized in production is presented. Such a discussion may prove useful in testing hypotheses formulated exclusively on the observation of emotional-prosodic defects in RHD patients (e.g., Ross, 1981; Tucker et al., 1977).

Linguistic Prosody

As noted earlier, the modulation of temporal and spectral parameters of the speech stream mark emotional attributes of the speaker and *additionally* serve various linguistic operations: signaling the modal function of speech acts (e.g., whether a statement or question is intended), lending prominence to material of specific importance within an utterance (emphasis), or preserving phonemic distinctions between lexical items of identical segmental structure. In response to claims that the right hemisphere may play a privileged role in the processing of affective prosody, Weintraub, Mesulam, and Kramer (1981) explored the ability of 9 RHD and 10 control subjects to express and discriminate *non-emotional* aspects of prosody. Production tasks required subjects to repeat sentences differing in intonation contour (statement-question) or in location of emphatic stress (e.g., *Steve drove the car* vs. *Steve drove the* ***car***). Additionally, subjects were required to produce emphasis in sentences following an elicitation paradigm (e.g., "*The boy ran down the alley*"; "*The girl walked down the street*"; *Who walked down the street?*). A single rater judged the accuracy of stress placement or the similarity of each recording to the model (where applicable).

Results indicated that the RHD patients were significantly impaired relative to the control subjects on *all* linguistic prosody tasks, including the production of emphatic stress and intonational contrasts. Coupled with previous data on affective prosody (Ross & Mesulam, 1979; Tucker et al., 1977), the authors hypothesized that the prosodic defect following right hemisphere insult may be more widespread than previously believed, extending to both its affective and linguistic components (Weintraub et al., 1981; also Shapiro & Danly, 1985). However, the generalizability of Weintraub et al.'s findings is restricted by the small number of test stimuli used and a potentially biased rating procedure.

Subsequent investigations of non-affective prosody in RHD patients also question the hypothesis that only emotional attributes of prosody are the province of the right hemisphere. As described in the preceding section, several studies have noted prosodic abnormalities in RHD patients when they were required to differentiate emotional *as well as* linguistic sentence types (Baum & Pell, 1997; Brådvik et al., 1991; Shapiro & Danly, 1985). Furthermore, in a recent acoustic comparison of prosody pre- and post-right CVA, Blonder and his colleagues (Blonder, Pickering, Heath, Smith & Butler, 1995) illustrated abberant prosodic patterning for non-affective stimuli in a RHD patient. Declarative utterances produced by the patient 6 months following her stroke demonstrated significantly "flattened" contours (i.e., little F_0 variation and lower F_0 peaks) relative to similar utterances recorded by the patient 6 months preceding the event. Collectively, these findings suggest that the RH's role in the production of prosody may be less tied to the domain of emotional expression than previously assumed (Baum & Pell, 1997; Blonder et al., 1995; Brådvik et al., 1991; Shapiro & Danly, 1985).

Cooper and his associates examined the production of non-affective intonation in small groups of RHD, LHD, and healthy adults using a reading task (Cooper, Soares, Nicol, Michelow & Goloskie, 1984). Sentences of varying lengths (e.g., *Al wants peaches*; *Al wants to buy some peaches*) were elicited from the subjects and the recordings were examined for differences in F_0 and timing attributes. Although statistical analyses were omitted from this investigation, trends in the acoustic data demonstrated abnormal use of speech timing and F_0 in the productions of the RHD patients relative to the control group, providing further evidence of right hemisphere involvement in the production of linguistic prosody (Cooper et al., 1984). Perhaps more importantly, the LHD aphasic patients in this study appeared to deviate from normalcy to a far *greater* extent than the RHD patients in the manipulation of both spectral and temporal parameters of sentence intonation. The possibility of left hemisphere involvement in the programming of non-affective prosody, at least at the sentence level, can therefore be inferred from these data (Cooper et al., 1984). This pattern of results, if corroborated, would be suggestive of incomplete lateralization or *bilateral* control of linguistic-prosodic functions.

Findings consistent with bilateral control of linguistic intonation have emerged elsewhere in the production literature. Ryalls, Joanette, and Feldman (1987) extracted acoustic measures of nonaffective speech elicited from 19 RHD and 9 control speakers of French: their analyses failed to uncover significant between-group differences in the use of F_0 or duration cues, prompting the authors to propose that linguistic prosody is not strongly lateralized in the brain (Ryalls et al., 1987). Moreover, timing or F_0-related acoustic anomalies for linguistic prosody have been reported in LHD aphasic patients in addition to RHD patients (Danly et al., 1983; Danly & Shapiro, 1982; Gandour, Holasuit Petty & Dardarananda, 1989). These data (indirectly) point to an allocation of resources *between* the two hemispheres for the production of linguistic prosody.

The production of non-affective intonation by RHD patients has also been investigated by Behrens (1989). To examine whether right hemisphere dysfunction affects the ability to manipulate specific F_0 properties of prosodic stimuli, she elicited utterances of different syntactic types (declaratives, imperatives, yes/no and WH-questions) from 8 RHD and 7 control subjects using a story completion task. Results indicated that the RHD patients produced contours that were generally normal in direction and rate of F_0 decline. Further, patients utilized changes in F_0 variance to distinguish three of the four sentence types in a manner ressembling that of the normal subjects (Behrens, 1989). These findings indicate that the RHD patients may have largely retained the ability to mark basic distinctions in propositional content.

Irregularities in the patterning of other aspects of intonation contours were evident in Behrens' RHD patients, however (e.g., the RHD patients produced a less linear declination slope than the normal subjects). The RHD patients may, therefore, have experienced difficulty in the "global" regulation of F_0 throughout the sentence (see also Baum & Pell, 1997). As the same RHD group proved capable of producing the F_0 correlates to local linguistic-prosodic forms such as contrastive and phonemic stress in a previous study (Behrens, 1988), the author hypothesized that the *domain* of prosodic expression, and not the communicative role of F_0 cues, may be central to the question of prosody lateralization (Behrens, 1989). Within such a framework, the right hemisphere may demonstrate superiority in encoding *large* domains of prosody such as the sentence or phrase, but is minimally involved in encoding smaller prosodic units at the word level (Behrens, 1989; see also Gandour et al., 1995).

Certainly, the study of how prosodic cues occupying relatively small domains (e.g., phonemic stress) are lateralized in production has resulted in far less disparity than the study of sentence-level prosodic phenomena. For instance, numerous studies have required unilaterally RHD subjects to disambiguate phonemic word pairs such as "**green**house" (house to grow plants) and "green house" (house that is green) in production; the results of these investigations have almost unanimously pointed to a sparing of this function in RHD patients, as determined by either perceptual ratings (Emmorey, 1987; Lebrun et al., 1985) or objective acoustic measures (Behrens, 1988; Emmorey, 1987; Hird & Kirsner, 1993; Ouellette & Baum, 1993).[1]

In one of the few studies to compare the production of phonemic stress by RHD and LHD aphasic patients, Emmorey (1987) reported that LHD nonfluent aphasics, but not RHD patients, were disturbed in the ability to distinguish noun compounds and noun phrases relative to matched control subjects. Taken together with the data indicating intact production of phonemic stress in RHD patients, Emmorey's (1987) findings argue strongly for a left hemisphere role in this linguistic function. Her acoustic measurements further revealed that none of the nonfluent aphasics used pitch to differentiate noun compounds from phrases, and only two nonfluent aphasics used duration. In contrast, all but one of the RHD and normal subjects employed pitch and/or duration cues in producing these distinctions. Based on this outcome, Emmorey (1987) postulated that the ability to produce pitch and duration cues may be dissociated at the lexical level, with duration cues showing greater resilience than pitch cues to left-hemisphere damage. However, this pattern of differential cue sensitivity was not observed in a follow-up study conducted by Ouellette and Baum (1993), as those authors were

---

[1]In the only exception to this pattern, Bryan (1989) reported impaired production of phonemic stress in RHD patients relative to LHD and normal patients. However, this assessment was based on a single rater's intuitions and may have been prone to error.

unable to detect significant differences between LHD, RHD and normal subjects in the production of phonemic stress, despite employing a highly comparable paradigm.

Support for the notion that lexically-assigned suprasegmental parameters are mediated by left hemisphere mechanisms is further garnered from investigations of *tone* production in brain-injured subjects. Independent studies of aphasic speakers of Norwegian (Ryalls & Reinvang, 1986), Mandarin (Packard, 1986), and Thai (Gandour et al., 1992) indicate that focal LHD often yields defects in the production of tonal contrasts when compared to matched control speakers. It is noteworthy that two of these investigations demonstrated a preservation of tonal production in comparable RHD groups when compared to the normal group (Gandour et al., 1992; Ryalls & Reinvang, 1986). Finally, the group of RHD Mandarin subjects reported to display difficulties in producing emotional-prosodic stimuli by Edmondson et al. (1987) were concurrently shown to exhibit intact production of phonemic tone. When considered collectively, the production data on phonemic tone present as relatively uniform, pointing to a left hemisphere superiority and minimal right hemisphere involvement in this function (Edmondson et al., 1987; Gandour et al., 1992; Packard, 1986; Ryalls & Reinvang, 1986). These data are also not inconsistent with the hypothesis that the size of prosodic units determine their lateralization in the brain (Behrens, 1989).

Far fewer researchers have examined the ability of brain-damaged subjects to produce *emphatic* stress, a skill requiring local manipulation of prosodic cues to highlight elements of value within a phrasal context. As outlined earlier, Weintraub et al.'s (1981) findings suggested that RHD patients may be impaired in repeating and spontaneously producing emphasis in short utterances. Similarly, Bryan's (1989) results indicated that both RHD and LHD patients produce less reliable cues to emphatic stress than normal subjects, and that RHD patients may be significantly more impaired than LHD patients. Unfortunately, the conclusions of both of these studies (Bryan, 1989; Weintraub et al., 1981) are weakened by a highly subjective rating procedure. In an acoustic investigation, Hird & Kirsner (1993) revealed temporal abnormalities in the production of contrastive items by RHD patients when compared to normal subjects. Interestingly, those authors attributed their findings to the possible pragmatic function of contrastive stress cues (i.e., marking the "givenness" or "newness" of information within a context), in light of the right hemisphere's recognized role in pragmatic functioning (e.g., Kaplan, Brownell, Jacobs & Gardner, 1990).

However, discordant findings have also emerged in the production literature on emphatic stress. Neither Behrens (1988) nor Ouellette and Baum (1993) uncovered acoustic evidence that RHD patients are disturbed in the production of emphatic stress, despite performing analyses on a far greater number of tokens than those examined in related studies (Bryan, 1989; Weintraub et al., 1981). Furthermore, Ouellette and Baum (1993) demonstrated that the capacity to encode emphasis would appear to be intact in LHD aphasic patients as well. These discrepancies in the literature, coupled with the relatively low number of investigations to address this issue, do not lead to definitive conclusions about the functional lateralization of emphatic stress production. This topic awaits further elucidation.

Summary of Production Data

As may be evident from the preceding review, much remains unclear about the neural representation of expressive prosodic functions. Although considerable research advocates a dominant (if not exclusive) role for the right hemisphere in the production of emotional-prosodic stimuli (Borod et al., 1985; Edmondson et al., 1987; Gandour et al., 1995; Hughes et al., 1983; Ross, 1981; Ross & Mesulam, 1979; Ross et al., 1981; Ross et al., 1988; Tucker et al., 1977), other reports suggest that the neural mechanisms underlying emotional prosody are not strongly lateralized or are distributed between the two hemispheres (Baum & Pell, 1997; Cancelliere & Kertesz, 1990). For the production of linguistic prosody, there is now considerable evidence that lexically-assigned prosodic cues of phonemic relevance are likely subserved by the intact left-hemisphere (Behrens, 1988; Edmondson et al., 1987; Emmorey, 1987; Gandour et al., 1992; Gandour et al., 1995; Hird & Kirsner, 1993; Lebrun et al., 1985; Ouellette & Baum, 1993; Ryalls & Reinvang, 1986). The neural substrates of our ability to encode prosodic features that occur over larger domains, such as contrastive stress or linguistic intonation, are poorly localized at present and await further inquiry.

Perception/Comprehension of Prosody

Perceptual investigations of prosody, like the production studies described in the foregoing discussion, have been largely concerned with how the functional significance of prosodic cues and/or their domain of processing influence the lateralization of prosodic stimuli in brain-injured adults. Perhaps distinct from the production literature, a greater proportion of the evidence for the perceptual lateralization of prosody has been derived from studies of *non-pathological* performance using the dichotic listening technique (Kimura, 1961). The relevant data arising from these studies are reviewed below.

Emotional Prosody

Studies of the perception and recognition of emotional attributes of speech have contributed greatly to the hypothesis that right hemisphere mechanisms are selectively engaged in the processing of affective prosody. In an early report that focused on the comprehension of affective speech, Heilman, Scholes, and Watson (1975) presented auditory stimuli to 6 LHD and 6 RHD subjects with temporoparietal lesions in two conditions: one in which subjects labelled the emotional mood of the speaker (happy, sad, angry, indifferent) and one in which subjects identified the semantic content of the same utterances. Judgments were indicated by pointing to line drawings of emotional facial expressions (emotion condition) or a graphic depiction of the semantic interpretation of the utterance (content condition) and the accuracy of each response was recorded.

Although both patient groups performed without error in interpreting the semantic meaning of the stimuli, results obtained in the emotion condition indicated that the RHD patients (who also presented with neglect) were significantly impaired relative to the LHD aphasic patients in the ability to categorize the affective meaning of prosodic cues, performing at near chance level. In a replication and extension of Heilman et al.'s (1975) study, Tucker, Watson, and Heilman (1977) obtained a similar pattern of results, reporting poorer comprehension of emotional prosody in RHD patients with neglect than in LHD aphasic patients (again, RHD patients identified the four emotions at chance level) . A disturbance in the ability to *discriminate* differences in prosodic patterns (i.e., make same/different judgements about paired stimuli differing in prosodic content) was also revealed by the RHD but not the LHD subjects (Tucker et al., 1977). Based on these data, both groups of investigators concluded that temporoparietal lesions of the non-dominant hemisphere in conjunction with neglect may lead to a selective impairment in the comprehension of affective prosody (Heilman, Scholes & Watson, 1975; Tucker et al., 1977). However, the strength of these assertions is limited by the observation that LHD patients were not error-free in their comprehension of affective meanings in either study, coupled with the absence of a non-neurological control group in both paradigms.

The results of more recent experiments have also been interpreted as supporting a privileged role for the right hemisphere in the receptive control of affective prosody. For example, Bowers, Coslett, Bauer, Speedie, and Heilman (1987) required 9 RHD, 8 LHD, and 8 healthy control subjects to identify emotional-prosodic meanings from several different types of stimuli, including utterances in which the semantic and prosodic message were either congruent or incongruent, and utterances that had been low-pass filtered of all identifiable linguistic content, preserving only the prosodic contour (i.e., "speech filtered" stimuli). For all the conditions tested, the RHD group exhibited significant deficits relative to both the LHD and control groups in the recognition of the emotional tone of the stimuli, suggestive of right hemisphere control of these processes. Similarly, Blonder, Bowers, and Heilman (1991) reported a global decline in the ability of their RHD patients to process the emotional significance of prosodic, facial, and gestural communicative signals when compared to LHD and non-neurological control subjects; this outcome was interpreted as indication of the primacy of the right hemisphere in the modulation of perhaps *all* (nonverbal) aspects of emotional communication.

To test the effects of "associational-cognitive" demands on the processing of emotional prosody, Tompkins and Flowers (1985) presented emotionally-intoned, semantically-neutral phrases to 11 RHD, 11 LHD, and 11 control subjects in three tasks of presumably increasing cognitive complexity: a discrimination task, an identification task in which subjects chose one of two possible emotional interpretations, and an identification task in which subjects judged the emotion from four possible alternatives. Consistent with other findings (Blonder et al., 1991; Bowers et al., 1987; Heilman et al., 1975; Tucker et al., 1977), the authors found that their RHD patients performed at an inferior level relative to the normal subjects on all emotional prosody tasks. However, the LHD patients' performance also broke down on the task in which the cognitive load was greatest (four-choice emotional identification). Thus, although their data were interpreted as further indication that the right hemisphere subserves emotional-prosodic processing, the authors postulated that the left hemisphere may become engaged as the cognitive demands of such tasks increase, as in tasks involving greater need for comparative processes or short-term memory (Tompkins & Flowers, 1985).

Evidence for right hemisphere superiority in recognizing emotional stimuli has also emerged from studies of normal prosody perception. Employing the dichotic listening paradigm, Ley and Bryden (1982) paired emotionally intoned (happy, sad, angry, neutral) and monotone sentences of similar grammatical construction for dichotic presentation to 32 young adults. Subjects were asked to attend to a specified ear and identify both the emotional tone and the verbal content of each sentence (independently for each ear) from a fixed set of alternatives. Analysis of the subjects' accuracy for each type of stimuli yielded a significant left-ear (right hemisphere) advantage for judging emotions and a significant right-ear (left hemisphere) advantage for judging the verbal content, with the majority of subjects (n=21/32) showing both trends simultaneously.

Thus, normative data indicating differential lateralization of emotional and verbal processing in the same subjects, consistent with a right hemisphere superiority in the comprehension of affective prosody, have come to light.

As noted earlier, Ross (1981) has not only advocated right hemisphere control of emotional prosody and gesture, but has elaborated a hypothetical model that places emotional-prosodic functions--both expressive and receptive--in circumscribed regions of the right hemisphere of the brain. In the receptive as well as the expressive mode, the work of Ross and his colleagues (Gorelick & Ross, 1987; Ross, 1981; Ross et al., 1981; Ross & Mesulam, 1979) has relied on bedside assessment of patients with acute right hemisphere lesions and suspected 'aprosodia'. To test affective comprehension, the examiner typically stands behind the patient and intones utterances with various affects (Ross, 1993; Ross et al., 1981). The patient is then asked to identify, either verbally or by means of a set list of alternatives, the emotion portrayed; in some cases, fewer than four out of five correct identifications has been considered impaired performance on this task (Gorelick & Ross, 1987). Employing this bedside technique, several case descriptions of RHD patients with posterior (temporoparietal) lesions and 'receptive aprosodia' (i.e., impaired affective comprehension in the face of spared affective production and repetition) have been described, each interpreted as validating Ross' proposed functional-anatomic organization of the aprosodias in the right hemisphere (Gorelick & Ross, 1987; Hughes, Chan & Su, 1983; Ross, 1981; Ross et al., 1981).

However, not all data are consistent with such a view. Individual cases of receptive aprosodia reported by several investigators (Brådvik et al., 1991; Darby, 1993; Heilman et al., 1984; Lebrun et al., 1985) clearly diverge from Ross' hypothetical model. Moreover, as described earlier, Cancelliere and Kertesz (1990) explored the relationship between acute vascular lesions and disturbances of emotional expression and comprehension and found no evidence that aprosodic deficits in RHD patients adhere to the anterior-posterior pattern described by Ross (1981); the authors attributed this discrepancy to their use of standardized stimuli and a less biased assessment procedure. Of even greater importance, Cancelliere and Kertesz reported emotional comprehension deficits of comparable frequency in both the right- *and* left-hemisphere-damaged adults they examined, calling into question the very notion that the right hemisphere is uniquely engaged in the processing of affective speech. Indeed, evidence that left hemisphere mechanisms also have some capacity to process emotional-prosodic stimuli would serve to explain purported cases of 'crossed aprosodia', or RHD patients without the predicted prosodic difficulties according to Ross' scheme (Gorelick & Ross, 1987).

Several investigations serve to corroborate Cancelliere and Kertesz' (1990) observations, providing evidence that both the right and left cerebral hemispheres contribute to the processing of affective vocal cues (Darby, 1993; Schlanger, Schlanger & Gerstman, 1976; Starkstein, Federoff, Price, Leiguarda & Robinson, 1994; Van Lancker & Sidtis, 1992). Using a picture-matching task, Schlanger,

Schlanger, and Gerstman (1976) presented emotionally-intoned utterances that were semantically neutral but meaningful (e.g., *He will come soon*) or semantically anomalous (*He will tuv roop*) to 20 RHD and 40 LHD aphasic patients for identification; these researchers found no significant differences in the accuracy of the two groups on this task, indicative of bilateral control of emotional prosody. Starkstein and his co-workers (1994) examined 59 consecutively-admitted patients with cerebrovascular lesions for the prevalence of receptive prosodic defects, conducting both neuropsychological and neuroradiologic analyses of these patients. Findings obtained for this study indicated that disturbed comprehension of affective prosody was a relatively frequent feature in both LHD *and* RHD acute stroke patients (45% of their sample), although RHD patients with basal ganglian or temporoparietal lesions exhibited a significantly greater incidence of such deficits. Thus, bilateral involvement in affective prosody comprehension is once again indicated, although Starkstein et al.'s (1994) findings further intimate the possibility that right hemisphere mechanisms may play a more predominant role.

A different perspective on the contributions of left and right hemisphere mechanisms in the comprehension of affective-prosodic stimuli was explored recently by Van Lancker and Sidtis (1992). They tested LHD, RHD, and healthy control subjects on an emotional prosody identification task and demonstrated no significant differences in the accuracy of the two clinical groups (which were both impaired relative to control subjects). To further explore whether the comprehension errors of LHD and RHD patients were predictable in terms of one or a combination of the acoustic parameters underlying emotional-prosodic meanings (i.e., to determine the extent to which either diagnostic group demonstrated impaired *perception* of the acoustic cues), the authors determined mean and variability measures of F_0, amplitude, and duration for the original stimuli. Discriminant function analyses were then performed to ascertain which of the acoustic cues served to signal the *intended* emotional meanings of the stimuli initially presented, and which cues predicted the comprehension *errors* made by each clinical group on the identification task; this was accomplished by recoding each emotional stimulus according to each group's most frequent error response for that stimulus. In this way, the authors sought to determine the extent to which the LHD and/or RHD subjects' emotional comprehension deficits were related to impaired perception of specific acoustic features of the stimuli.

Despite the similar level of impairment of LHD and RHD patients in identifying affective-prosodic meanings, analyses performed on each group's identification errors suggested that LHD and RHD patients were using the acoustic cues to prosody differently in judging affective meanings (Van Lancker & Sidtis, 1992). Interestingly, the discriminant analysis of the LHD subjects' errors revealed that these patients may have been basing their decisions on fundamental frequency information (particularly F_0 variability), whereas an analysis of the RHD subjects' affective misclassifications indicated a reliance on durational cues in identifying the stimuli. This pattern of results suggested to the authors that receptive disturbances of emotional prosody may be perceptual in nature, possibly reflecting the superiority of each hemisphere for the processing of different acoustic cues to prosodic meanings (Van Lancker & Sidtis, 1992). More generally, the authors concluded that, in contrast to previous proposals (e.g., Ross, 1981), the comprehension of prosody is a multifaceted process subserved by distributed (i.e., bilateral) mechanisms that are not strictly localizable to the right hemisphere.

The notion that brain-damaged patients may have a more basic disturbance in analyzing the acoustic structure of prosody is consonant with reports that these patients often show deficits in using the same auditory cues in *nonlinguistic* tasks. More specifically, RHD individuals have frequently been noted to make errors on nonlinguistic tasks that require the processing of complex pitch information, indicating that this skill may rely predominantly on right-hemisphere auditory mechanisms (Robin, Tranel & Damasio, 1990; Sidtis & Feldmann, 1990; Zatorre, 1988; Zatorre, Evans & Meyer, 1994). Interestingly, a left-hemisphere bias has been proposed for the processing of *temporal* cues on similar nonlinguistic tasks (Carmon & Nachshon, 1971; Robin et al., 1990). Collectively, these data are consistent with the interpretation that each hemisphere may contribute independent auditory processing capabilities to the task of decoding emotional stimuli (Van Lancker & Sidtis, 1992). However, in an attempt to replicate Van Lancker and Sidtis' preliminary findings, Pell and Baum (1997b) found no evidence that the emotional comprehension errors committed by their LHD and

RHD patients were biased by specific acoustic features of the stimuli, despite careful adherence to the authors' original methods. Thus, although intriguing, the hypothesis that individual acoustic cues to prosody are independently lateralized (Van Lancker & Sidtis, 1992) remains speculative and awaits future elucidation.

Linguistic Prosody

Thus far, our consideration of theories of receptive prosodic lateralization has concentrated on affective prosody, but the discussion may benefit from a review of the *linguistic* functions of prosodic cues as well. As noted in the preceding section, prosodic features expressed over various domains signal differences in the illocutionary intent of an utterance (e.g., whether information is stated or requested), highlight items of relative importance in a spoken message (emphasis), or disambiguate the meaning of words with similar segmental structure (phonemic stress). Several investigators have explored the neural basis for comprehension of *locally* defined linguistic-prosodic features such as phonemic or emphatic stress. In response to contentions in the literature that right hemisphere lesions selectively disrupt affective prosody, Weintraub, Mesulam, and Kramer (1981) tested 9 RHD and 10 control subjects for the comprehension, production, and repetition of linguistic prosody. One receptive task measured subjects' accuracy in discriminating phonemic stress contrasts (e.g., **green**house vs. green **house**) using a picture-identification paradigm and another measured their accuracy in making same/different judgments about sentence pairs differing in emphatic stress location (e.g., **Steve** drives the car vs. Steve drives the **car**) or intonation contour (statement vs. question).

Results obtained for each linguistic prosody task revealed significant impairments in the RHD group relative to the control subjects, a pattern interpreted as evidence that the RH's role in prosody may extend beyond its affective components to the linguistic domain (Weintraub et al., 1981). More recently, Brådvik, Dravins, Holtås, Rosén, Ryding, and Ingvar (1991) compared the performance of 20 Swedish-speaking patients with stable right hemisphere lesions and 18 normal controls on tasks of both linguistic and affective prosody (e.g., emphatic stress perception, identification of linguistic and emotional intonation) and arrived at a quite similar conclusion: the inferior performance of their RHD patients on both linguistic and emotional tasks pointed to an essential role for the right hemisphere in the processing of *both* (linguistic and affective) prosodic "functions", irrespective of the domain over which prosodic cues were perceived. The potential relationship between subcortical infarcts and a lasting disturbance of speech prosody, alluded to in the discussion of expressive prosodic deficits (Cancelliere & Kertesz, 1990), was also highlighted by their data (Brådvik et al., 1991).

The omission of a comparable LHD patient group in the latter two studies (Weintraub et al., 1981; Brådvik et al., 1991) again unfortunately impedes an appropriate understanding of each hemisphere's potential involvement in prosodic perception. In a study that considered RHD, LHD, and non-neurological control subjects simultaneously (n=30/group), Bryan (1989) presented a battery of linguistic prosody tests that incorporated stimuli of various perceptual domains (e.g., phonemic/emphatic stress discrimination, identification of declarative vs. interrogative intonation). Bryan demonstrated that the RHD patients were impaired on receptive tasks of linguistic prosody relative to both normal (13/13 tasks) and LHD (8/13 tasks) subjects, again favouring a right hemisphere basis for this processing. However, it is noteworthy that the LHD group reported by Bryan (1989) was significantly impaired relative to the control group on 10 of the 13 tests as well, a finding the author conceded may be suggestive of bilateral control for at least some aspects of linguistic prosody. This pattern of results may be conducive to a superior (albeit not exclusive) role for the right hemisphere in the comprehension of linguistic prosody, along the lines suggested earlier for the comprehension of emotional prosody (Starkstein et al., 1994).

However, still more research has placed the receptive control of linguistic prosody--at least, the perception of locally-assigned stress cues--firmly in the left hemisphere of the brain. For instance, Baum, Daniloff, Daniloff, and Lewis (1982), following Blumstein and Goodglass (1972), presented three tasks of stress comprehension to 8 LHD nonfluent aphasics and 8 control subjects and reported a significantly reduced capacity to comprehend phonemic and emphatic stress in their LHD patients, findings inconsistent with the notion that linguistic prosody is processed solely by the right hemisphere. Emmorey (1987) presented phonemic stress pairs (e.g., **hot**dog , hot **dog**) to 7 RHD, 15 LHD, and 22 control subjects for perceptual recognition and observed a significant decrement in the performance of the LHD subjects (both fluent and nonfluent aphasics) relative to control subjects on this task, but intact comprehension of the stimuli by RHD patients. These data corroborate and extend those of Baum et al. (1982), indicating a left hemisphere substrate for the perception of linguistic stress.

Related findings are derived from Behrens (1985) in which the dichotic listening technique was used; she required 15 normal subjects to identify stress placement in phonemic stress pairs and demonstrated a significant right-ear (left hemisphere) advantage on this task. Filtering the same stimuli at 200 Hz for presentation or reducing the semantic content of the stimuli (e.g., **botgog**) did not lead to a right-ear advantage, however, suggesting to the author that left hemisphere mechanisms process stress contrasts *except* when those cues are of minimal linguistic import (as in the speech-filtered stimuli). The results of these studies (Baum et al., 1982; Behrens, 1985; Emmorey, 1987) may be viewed as support for the "functional load" hypothesis of prosodic lateralization, or the notion that the linguistic or emotion role of prosodic cues in speech determines the laterality of processing (Van Lancker, 1980).

Further evidence that the left hemisphere may underlie our ability to perceive local, linguistically-assigned prosodic features is gleaned from perceptual investigations of languages in which pitch contrasts serve as a phonemic marker (e.g., Mandarin, Thai). The outcome of dichotic studies with normals (Van Lancker & Fromkin, 1973) and lesion investigations (Gandour & Dardarananda, 1983; Hughes et al., 1983) is in general agreement, demonstrating a left hemisphere bias for the ability to discriminate tonal distinctions by native tone-language speakers. These reports, in conjunction with the English data reviewed above (Baum et al., 1982; Behrens, 1985; Emmorey, 1987), provide substantial evidence to suggest a privileged role for the left hemisphere in the processing of linguistically-assigned prosodic cues expressed at the segmental or syllabic level.

However, affective prosodic features are typically expressed and perceived over domains *larger* than the word, usually the phrase or utterance. It is at this level of segmental structure--the sentence--that receptive studies of affective prosody have largely focused, and accordingly, that we are best able to compare the perception of prosodic cues as an index of their (linguistic or affective) "functional load" in speech. Regrettably, the perceptual literature on linguistic intonation is relatively small when compared to that on emotional intonation. In an early study to consider linguistic sentence prosody, Blumstein and Cooper (1974) presented dichotically paired utterances differing in intonational content to 40 young adults. In two separate experiments, subjects identified speech-filtered exemplars of the dichotic stimuli by their intonational meaning (declarative, interrogative, imperative, conditional) or matched the intonation pattern of filtered or nonsense (e.g., *padaka*) dichotic stimuli with a successively-presented foil. The accuracy of the subjects was then analyzed to determine the presence of an ear advantage on each task.

In general, the results revealed a "small but consistent" left ear (right hemisphere) advantage for all tasks of perceiving and identifying linguistic intonation. The authors concluded from their findings that linguistic prosody, in the absence of meaningful segmental (i.e., semantic) structure, may be processed more efficiently by the right hemisphere, and that even when recognizable segmental information is present in the stimuli (as was the case for the nonsense stimuli), left hemisphere mechanisms are likely minimally implicated at best (Blumstein & Cooper, 1974). These findings were contrary to the authors' original expectations that linguistic intonation may be processed by the left hemisphere in a way similar to other linguistic systems; the results provide tentative support to those investigators who have posited superior right hemisphere processing of sentence prosody generally, regardless of its function (Brådvik et al., 1991; Weintraub et al., 1981; but review Cancelliere & Kertesz, 1990; Darby, 1993; Schlanger et al., 1976; Starkstein et al., 1994; Van Lancker & Sidtis, 1992, for data indicating left hemisphere control of emotional sentence prosody).

Few studies have attempted to explore how each hemisphere is specialized to process sentence prosody in *both* linguistic and affective contexts concurrently. In one such study, Heilman, Bowers, Speedie, and Coslett (1984) presented auditory stimuli to 8 RHD, 9 LHD, and 15 control subjects in two identification tasks, one in which intonation conveyed the linguistic modality of the utterance (declarative, interrogative, imperative) and another in which prosodic cues signaled various affective meanings (angry, sad, happy). Stimuli in both conditions were speech-filtered before presentation, rendering the segmental content (but not the prosodic contour) unintelligible to the listener. Subjects indicated their response either verbally or by matching the prosodic meaning with an appropriate graphic representation (facial expression or punctuation mark for the affective and linguistic stimuli, respectively) and the accuracy of each response was measured.

Results obtained for this investigation revealed that the RHD patients made significantly more errors than both the LHD and control subjects (who also differed significantly) in identifying the emotional meaning of speech-filtered utterances, whereas the RHD and LHD patients were equally impaired relative to normals in identifying the linguistic intent of the stimuli. Moreover, only the LHD subjects' comprehension of prosodic meanings was affected by the *type* of prosody tested; specifically, the LHD subjects performed at a significantly inferior level on the linguistic task when compared to the emotional task, a pattern not observed for either the RHD or control groups. To account for these results, two hypothetical explanations were proffered (Heilman et al., 1984). Firstly, the processing of affective prosody may be lateralized to the right hemisphere of the brain (RHD patients were most impaired on this task) whereas the processing of linguistic intonation may be achieved bilaterally (both patient groups were impaired relative to normals). Alternatively, the right hemisphere may dominate *all* processing of sentence intonation (both linguistic and affective), but the left hemisphere becomes engaged on tasks as the need for linguistic processing increases (i.e., on non-affective tasks). The authors acknowledged that their data allowed for either interpretation; however, it is noteworthy that both proposals are inconsistent with previous assertions that the right hemisphere is specialized to decode only the *affective* features of prosodic stimuli (Blonder et al., 1991; Bowers et al., 1987; Ehlers & Dalby, 1987; Heilman et al., 1975; Ley & Bryden, 1982; Ross et al., 1981; Tucker et al., 1977).

In a recent investigation, Pell & Baum (1997a) administered identification tasks for both affective- and linguistic-prosodic stimuli to 9 RHD, 10 LHD, and 10 control subjects, testing the same target meanings employed by Heilman and colleagues (1984). To additionally address inconsistencies in the literature on prosody with respect to the type of stimuli presented (e.g., filtered, natural), linguistic and emotional stimuli were each presented in three distinct identification tasks: a semantically "well-formed" condition, in which both prosodic and semantic information cued the (linguistic or affective) intonational target meaning; a "nonsense" condition, in which phonetically-plausible but meaningless utterances were intoned to convey prosodic meanings corresponding to those presented in the well-formed stimuli; and, a "filtered" condition, in which the well-formed utterances were low-pass filtered to obscure the linguistic content but retain prosodic cues. A task requiring subjects to make same/different judgments about pairs of speech-filtered utterances was also presented to test for an underlying perceptual deficit in the subjects' ability to process prosodic information (Van Lancker & Sidtis, 1992). Both accuracy and response time data were collected.

Although all three groups were shown to perform comparably in discriminating prosodic patterns, results of the identification tasks revealed that neither the LHD nor RHD subjects were impaired relative to normals in recognizing the emotional meaning of prosodic patterns (angry, sad, happy), but that both clinical groups exhibited deficient comprehension of linguistic-prosodic meanings (declarative, interrogative, imperative). Interestingly, comparing the performance of each group across linguistic and affective domains revealed a pattern qualitatively similar to that reported by Heilman et al. (1984); RHD and control subjects each demonstrated similar capabilities on corresponding linguistic and affective tasks (reflected in both their accuracy and response times), whereas LHD aphasic subjects *always* responded significantly slower and with less precision on the linguistic relative to the affective task (even though no semantic comprehension of the stimuli was required).

Thus, a specific susceptibility to the linguistic load of prosodic stimuli was again noted in LHD but not RHD adults (Heilman et al., 1984; Tompkins & Flowers, 1985), although it is important to bear in mind that RHD patients were also impaired for the linguistic stimuli. The perhaps surprising observation that neither clinical group was impaired in the comprehension of emotional prosody may have been due to clinical differences between Pell & Baum's patients, who had been screened for behavioural neglect, and those tested elsewhere; indeed, the coincidence of lasting aprosodias and severe neurologic signs such as neglect have been noted previously on several occasions (Heilman et al., 1975; Starkstein et al., 1994; Tucker et al., 1977). Overall, the results of this study highlight the possibility that receptive prosodic functions, both linguistic and affective, may not be subserved by mechanisms lateralized to a single hemisphere of the brain, at least not when this processing occurs over larger domains such as the sentence (Pell & Baum, 1997a).

Finally, other perceptual research suggests that the locus of mechanisms subserving prosody may not be limited to cortical regions, but rather, may be organized subcortically. In particular, the basal ganglia have been implicated as a structure of potential importance in several investigations of vascular patients with receptive aprosodias reviewed above (Brådvik et al., 1991; Cancelliere & Kertesz, 1990; Ross & Mesulam, 1979; Starkstein et al., 1994). These findings obtain further support from studies that have examined receptive prosody in patients with basal ganglia dysfunction as a result of Parkinson's or Huntington's disease (Blonder, Gur & Gur, 1989; Borod et al., 1990; Cancelliere & Hausdorf, 1988; Pell, 1996; Scott, Caird & Williams, 1984; Speedie, Brake, Folstein, Bowers & Heilman, 1990). For example, Blonder, Gur, Gur (1989), and more recently Pell (1996), each demonstrated impaired comprehension of linguistic and emotional intonation in idiopathic Parkinsonian patients relative to healthy control subjects; coupled with the cortical data on receptive prosody, the outcome of each of these investigations would appear to advocate a functional network dedicated to prosody consisting of both cortical and subcortical components (Blonder et al., 1989; Pell, 1996). The issue of subcortical representation of prosody in both receptive and expressive behaviour is therefore worthy of pursual in future investigations.

Summary of Perception Data

Our review of receptive investigations of prosody converges with that of production studies in its (weak) support of differential lateralization of prosodic cues as an index of their (linguistic or affective) communicative function in speech. To date, results emanating from studies of phonemic stress and pitch perception have demonstrated relatively consistent involvement of the left hemisphere and relatively infrequent involvement of the right hemisphere, signifying a left hemisphere neural substrate for linguistically-relevant prosodic cues operating over short domains (Behrens, 1985; Emmorey, 1987; Van Lancker, 1980). However, it is at the sentential level that the effects of the functional load of prosodic cues become more opaque, and the issue of laterality becomes less certain. Although ample evidence has now accrued to suggest that the affective attributes of prosody are not processed uniquely by the right hemisphere (Cancelliere & Kertesz, 1990; Darby, 1993; Dykstra, Gandour, & Stark, 1995; Heilman et al., 1984; Pell & Baum, 1997a; Schlanger et al., 1976; Seron et al., 1982; Tompkins & Flowers, 1985; Van Lancker & Sidtis, 1992), it remains unclear as to whether the right hemisphere serves a dominant (albeit shared) role in the processing of emotional and linguistic prosody (Blumstein & Cooper, 1974; Heilman et al., 1984; Starkstein et al., 1994) or whether emotional and linguistic prosody functions are distributed bilaterally (Bryans, 1989; Cancelliere & Kertesz, 1990; Dykstra, Gandour & Stark, 1995; Pell & Baum, 1997a; Van Lancker & Sidtis, 1992. Finally, subcortical structures may be critical in the regulation of prosodic functions in receptive and expressive modalities (Brådvik et al., 1991; Blonder et al., 1989; Cancelliere & Kertesz, 1990; Pell, 1996).

Present Objectives

As may be gleaned from the foregoing review, right hemisphere participation in prosodic functions is suggested by certain trends in the production and perception literatures on prosody. Most notably, emotional-prosodic features are impaired most commonly in patients with right hemisphere insult, whereas locally-assigned linguistic-prosodic features (e.g., phonemic stress and tone) tend to be disturbed relatively infrequently in RHD populations. Unfortunately, this general pattern in the prosody data fails to exemplify what communcative parameters specify the right hemisphere's role in modulating prosodic stimuli; for example, such a pattern does not illuminate the extent to which functional attributes, the domain of processing, or specific acoustic properties of prosodic stimuli are related to right hemisphere specialization for prosodic functions. Methodological protocols implemented to date have been largely unsuccessful in delineating the contribution of these various factors to right hemisphere involvement in prosodic processing.

Investigating the manner in which focal right-brain injury affects prosody when multiple cues--both linguistic and affective in nature, and encompassing different processing domains--are manipulated concurrently by the same subjects may begin to serve this utility. Such an approach could yield considerable insight into the biological basis of prosodic functions, and help reconcile some of the factors underlying the cerebral lateralization of prosody encoding and decoding.

To explore the interaction of linguistic and emotional cues to prosody in normal speech production and perception, and to illuminate the underlying acoustic basis of prosodic deficits in right-brain-damaged adults, two experiments were designed. In Experiment 1, utterances of two distinct lengths ("short" and "long") are elicited from RHD and normal control (NC) subjects to investigate the concomitant effects of sentence focus, sentence modality, and emotional intent on various acoustic properties of the stimuli. In Experiment 2, the RHD and control subjects tested in Experiment 1, as well as a control group of unilaterally left-hemisphere-damaged subjects, are presented six receptive tasks; three tasks require subjects to locate the position of emphatic stress within an utterance, and three tasks require subjects to identify a speaker's emotional state from the prosody. Perceptual tasks employ stimuli ressembling those generated in Experiment 1, but in which either the F_0 or duration parameters of the stimuli are "neutralized" by means of acoustic manipulation prior to the experiment. This process allows the contribution of each acoustic parameter in the perception of focus and emotion to be investigated, for each of the three subject groups.

Through these experiments, a number of hypotheses advanced previously in the literature on prosody may be scrutinized. First, by requiring RHD patients to encode and decode the linguistic and emotional features of prosodic stimuli *in tandem*, the hypothesis that only emotional aspects of prosody are modulated by the right cerebral hemisphere may be tested (e.g., Blonder et al., 1991; Heilman et al., 1975; Ross, 1981; Tucker et al., 1977). In production (Experiment 1), this hypothesis would predict acoustic anomalies in the speech of RHD patients, but not controls, for those parameters important in signaling emotional distinctions (e.g., mean F_0, F_0 range, speech rate). However, no between-group differences should emerge in the patterning of linguistically-relevant acoustic contrasts (e.g., keyword F_0 and duration, utterance terminal, focus accent).

In perception (Experiment 2), corroboration of the "functional role hypothesis" would be indicated by a significant decrement in recognition performance across the three emotion tasks relative to the three focus (linguistic) tasks for the RHD group. In addition, the RHD patients should demonstrate significantly reduced accuracy on the emotion tasks, overall, relative to both the LHD and normal subjects. Failure to uncover a significant group effect may be interpreted as evidence that less lateralized processes underlie emotional prosody decoding (Cancelliere & Kertesz, 1990; Pell & Baum, 1997a; Van Lancker & Sidtis, 1992).

The current protocol simultaneously explores how prosodic constituents of different domains influence right hemisphere processing of the stimuli. This follows suggestions that "global" prosodic features, occurring over relatively large domains (such as the sentence), may be more conducive to right hemisphere processing mechanisms than locally-defined prosodic features (Behrens, 1989; Emmorey, 1987; Gandour et al., 1992). In Experiment 1 (Prosody Production), the manner in which this hypothesis was tested was two-fold: subjects were required to transmit prosodic messages associated with different levels of linguistic structure (focus is assigned to individual words, whereas linguistic modality and emotion are planned over the utterance as a whole); and, subjects were required to modulate prosodic features in utterances varying in overall length (6 or 10 syllables).

According to the "domain hypothesis", the RHD patients should present with difficulties regulating the acoustic attributes to both linguistic and emotional intonation, demonstrating aberrant production of (for example) mean F_0, F_0 range, speech rate, and/or utterance terminal measures. By contrast, the acoustic underpinnings to emphatic stress (focus accent, keyword F_0 and duration) should demonstrate a relatively normal distribution in the RHD patients. Furthermore, acoustic irregularities may prove more abundant in the RHD patients' production of "long" versus "short" utterances, although such an effect has not previously been established in the production literature on prosody.

In Experiment 2 (Prosody Perception), the domain hypothesis should yield the same pattern of performance as that predicted by the functional role hypothesis: impaired recognition of emotional cues relative to emphatic cues for the RHD patients, but the opposite pattern (impaired focus but not emotion perception) for the LHD patients. The effects of overall sentence length on receptive prosody performance is not investigated.

Hemispheric specialization for the modulation of specific acoustic parameters of prosodic stimuli has also been proposed (Dykstra et al., 1995; Emmorey, 1987; Robin et al., 1990; Van Lancker & Sidtis, 1992). In particular, the right hemisphere may assume a privileged role in controlling F_0 attributes of prosodic stimuli (Robin et al., 1990; Van Lancker & Sidtis, 1992; cf. Pell & Baum, 1997b). In Experiment, 1, this possibility is investigated by examining what types of acoustic measures (if any) demonstrate a significant departure from normalcy in the RHD patients' productions; a defect in the implementation of F_0, but not duration, at *all* levels of prosodic structure (focus, modality, emotion) may be interpreted in light of the "cue lateralization hypothesis". Such a pattern would indicate substantial (although not necessarily exclusive) right-hemisphere involvement in the regulation of F_0, and point to a dissociation of the neural mechanisms underlying F_0 and duration encoding for prosody.

For receptive abilities, the manipulation of specific F_0 or duration characteristics of the stimuli presented in Experiment 2 permits a direct test of the cue lateralization hypothesis. Since stimuli were modified in such a way as to "neutralize" the influence of either F_0 or duration relative to a baseline task, the strength of each cue in recognizing focus and emotion could be established for each of the three subject groups. Given the hypothesis that RHD patients are disturbed in the perception of F_0 attributes of prosodic stimuli (Robin et al., 1990; Van Lancker & Sidtis, 1992), one would anticipate inferior performance on the part of the RHD group on tasks in which *only* F_0 cues suggest a possible response (i.e., focus and emotion tasks in which duration cues were neutralized). The LHD patients may exhibit the opposite pattern, whereby the *removal* of F_0 cues leads to differential impairment in prosody recognition (Robin et al., 1990; Van Lancker & Sidtis, 1992). Such deficits, if observed, should occur irrespective of the linguistic or emotional nature of the stimuli, or the size of the prosodic unit being processed.

Thus, more directly and comprehensively than in previous investigations, the experiments reported herein seek to assess how RHD patients produce and identify various aspects of prosodic stimuli. In this manner, it may become apparent which elements of prosodic structure (if any) pose specific difficulties for RHD patients to produce, and which cues inhibit or foster the patients' comprehension of the stimuli. Illuminating some of the demands involved in expressing and perceiving multiple prosodic distinctions in RHD individuals may facilitate a clearer definition of the right hemisphere's role in regulating prosodic functions, and prove useful in informing current theories of prosody lateralization (Behrens, 1989; Ross, 1981; Van Lancker & Sidtis, 1992).

Finally, invaluable normative acoustic data reflecting the superimposing of influences of linguistic and emotional prosody may be garnered from this inquiry. These data will be useful in corroborating previous acoustic descriptions of the interaction between emphatic stress and sentence-type contrasts (Cooper et al., 1985; Eady & Cooper, 1986; Ryalls et al., 1994), and more importantly, will *extend* this knowledge to various emotional contexts. This preliminary work may prove constructive in refining our current understanding of the vocal correlates to discrete emotions, and suggest how emotional parameters are systematically altered in sentences where linguistic-prosodic contrasts must simultaneously be preserved. Ultimately, it is hoped that through the present data, an acoustic description of the prosodic modifications that take place in natural discourse may begin to emerge.

EXPERIMENT 1

To explore how linguistic and emotional prosody interact in the acoustic signal, and to reconcile the extent to which various aspects of prosodic structure implicate right hemisphere encoding mechanisms, the production of prosody was studied in unilateral RHD patients and matched subjects without neurological dysfunction. Experiment 1 sought to provide preliminary acoustic data reflecting the ability of RHD patients, relative to that of age-matched control subjects, to express prosodic distinctions in utterances where the underlying acoustic parameters varied in function (linguistic vs. affective), domain (word vs. utterance; short vs. long utterances), and/or degree (i.e., the number of distinctions necessary to manipulate at one time).

METHODS

Subjects

Ten (10) subjects with unilateral right-hemisphere-damage (X=64.3, range=29-87), and 10 healthy control subjects without neurological dysfunction (X=66.1, range=59-72) volunteered to take part in the study. Clinical subjects were recruited from hospitals in the Montreal (Quebec) and Cornwall (Ontario) regions. Control subjects represented members of an active database of individuals in the greater Montreal region willing to participate in language and memory studies at the School of Communication Sciences and Disorders at McGill University.

All subjects were native English speakers and all but one (R8) were right-handed. Apart from the single-event CVA suffered by RHD patients, no subject evidenced signs of neurological or psychiatric illness prior to the study, as determined from medical records (RHD patients) or by questionnaire (NC subjects). Puretone air conduction screenings of both ears at 0.5, 1, and 2 kHz ensured that all subjects had acceptable hearing levels before participating in the experiment; inclusion criteria were set at 30 dB HL at each frequency, for the better ear. The RHD and NC groups were closely balanced for gender. Basic demographic and clinical attributes of the RHD patients are furnished in Table 1.

For each RHD subject, CT scan confirmed the presence and location of the offending lesion within the right-hemisphere. None of the clinical subjects presented with aphasic deficits subsequent to right hemisphere CVA. Hemispatial visual neglect was identified in four (4) RHD patients, as assessed using the Bells Test (Gauthier, Dehaut & Joanette, 1989). Other behavioural 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. Finally, it should be noted that patients were tested at least three months post-onset of stroke (X=36

Table 1. Basic demographic and clinical characteristics of the RHD subjects participating in Experiment 1.

Sbj	Sex	Age	Post-Onset	Lesion Site	Major Clinical Signs	VN
R1	F	54	66	(R) posterior	(L) hemiparesis, flat affect	–
R2	M	59	31	(R) temporo-parietal	(L) hemiparesis, inappropriate mood, impaired inferencing/ comp. of figurative language	+
R3	M	74	48	(R) parietal	(L) hemiparesis, impaired comp. of figurative language	–
R4	F	61	8	(R) basal ganglia	(L) hemiparesis, flat affect	–
R5	F	29	7	(R) MCA	(L) hemiplegia	–
R6	F	82	40	(R) temporal	(L) hemiplegia, impaired recognition of emotional prosody	+
R7	M	69	59	(R) temporo-parietal	(L) hemiparesis, impaired recognition of emotional prosody	–
R8	F	87	83	(R) MCA	(L) hemiparesis, impaired inferencing/comp. of figurative language	–
R9	F	62	10	(R) external capsule	(L) hemiparesis, flat affect, impaired inferencing/comp. of figurative language/emotional prosody recognition	+
R10	F	66	12	(R) fronto-parietal	(L) hemiparesis, impaired comp. of figurative language	+

Note: Age=years, Post-Onset=months, VN=visual neglect for the left hemispace.

months, range=7-83 months) and therefore exhibited a relatively stable clinical profile at the time of testing.

Stimuli

Four sentences, two of which were 6 syllables in length ("*short*") and two of which were 10 syllables in length ("*long*"), served as the test stimuli. Each sentence was constructed so that it could be read without emphasis, or with emphasis in one of two "keyword" positions: sentence-initial or sentence-final.

<u>*Short*</u>	<u>*Long*</u>
Barry took the *sailboat*	*Barry* took the sailboat for the *weekend*
Mary sold the *teapot*	*Mary* sold the teapot for a *dollar*

As may be seen, keywords (italicized) constituted frequently observed content words matched across the stimuli for syllable length and stress location at each keyword position. All stimuli followed the canonical subject-verb-object syntactic ordering of English utterances. "Long" items were distinguished from "short" items by a terminal prepositional phrase.

In addition to varying the location of focus within each sentence, care was taken in stimulus preparation that each utterance could be intoned as both a statement and a yes-no question (without subject-auxillary inversion). Moreover, the content of each sentence was conducive to four different affective readings (sorrow, joy, anger, and neutral). Thus, 24 versions of each of the four test stimuli--each utterance representing a unique combination of the three prosodic variables (focus, sentence modality, emotion)--were elicited, for a total of 96 productions per speaker (4 items X 3 focus locations X 2 modalities X 4 emotions). To facilitate the manipulation of prosodic cues *without* altering the influence of segmental structure between conditions, "neutral" semantic content that did not suggest a particular linguistic or affective interpretation was adopted. The content of a pre-recorded priming stimulus preceding each trial served as the vehicle by which specific combinations of prosodic attributes were elicited (Cooper et al., 1985; Eady & Cooper, 1986).

To systematically bias the subjects' reading of the stimuli to obtain specific constellations of prosodic features differing in focus location, sentence modality, and emotion, short passages or "scenarios" were constructed. For example, the scenario preceding a nonemotional, declarative reading of *Mary sold the teapot* lacking 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:*

To elicit focus at one of the two keyword positions, the scenarios were modified to ensure that the information to be used contrastively was not "given" in the prime. When emotional readings of the stimuli were sought, scenarios provided an explicit context in which the speaker was described as being happy, sad, or angry before listening to a prompt similar to that used in "neutral" contexts. Cues cards containing the target utterance and information reinforcing how the sentence was supposed to be read (e.g., for focus, the target word was indicated in bold italics) further facilitated the desired reading for each trial.

Priming scenarios were recorded in a sound-proof 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 relatively neutral, "reporting" tone that did not lend undue prominence to particular lexical items within the passage. The 96 scenarios were subsequently randomized for order of presentation and the tape was edited to reflect this random order. A five second 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 for the RHD patients and in a laboratory setting for the control subjects. To ensure that subjects were attending closely to the content of the priming scenarios, testing was completed during two separate 30 minute sessions, half of the trials (n=48) being presented during each visit.

Subjects were seated comfortably at a table with a directional microphone (Sony ECM-909) placed 20 centimeters in front of their mouths. A small binder containing the stimulus cards corresponding to each trial was placed on the table in front of each subject. Prior to testing, subjects were encouraged to listen closely to the priming "stories" and then read what was written on the card in front of them; priming contexts were presented free-field to each subject using a Sony portable radio-cassette recorder. When reading errors or dysfluencies occurred during the experiment, subjects were requested to repeat the sentence. As well, subjects were informed that they were free to repeat their responses at any time during the experiment if they were not completely satisfied with their performance, although it is noteworthy that such "corrections" were observed infrequently.

Whenever multiple productions occurred for the same trial, the final token produced was always considered in subsequent analyses. Five practice trials preceding the experiment helped subjects become acquainted with the procedure, and although subjects occasionally commented that priming scenarios were somewhat repetitious, they demonstrated little difficulty in performing the task. All responses were recorded by a Sony (TCD-D3) digital audio tape recorder for later analysis.

Acoustic Analyses

Subjects' productions were digitized using the BLISS speech analysis system (Mertus, 1989) at a sampling rate of 20 kHz, with a 9 kHz low-pass filter setting and 12-bit quantization. The 96 utterances generated for each speaker were then acoustically analyzed using the BLISS waveform editor. To appropriately characterize the linguistic- and emotional-prosodic distinctions of interest herein, F_0 and duration measures were extracted at a number of points in each utterance. As different aspects of prosodic structure tend to be associated with acoustic alterations over various domains (e.g., the acoustic manifestation of focus tends to be more localized within the speech stream than that of emotional content), both "local" and "global" fluctuations in F_0 and duration were determined for each utterance. In all cases, acoustic measures were determined on the vowel segment at the point of interest, as consonants have been shown to be relatively unaffected by prosodic highlighting (Brown & McGlone, 1974; Fry, 1955; Klatt, 1976). Amplitude measures were omitted from the present analyses, in light of the minimal significance accorded to these cues in several previous reports (Behrens, 1988; Bolinger, 1958; Morton & Jassem, 1965; Ross et al., 1988; Turk & Sawusch, 1996). For each utterance produced, the following acoustic measures were derived:

"Local" Measures

1) *Keyword Duration* - to characterize temporal changes throughout the utterance, the duration of the full vowel (transition and steady state) was determined on the stressed syllable of each content word. Through both visual and auditory analysis of the oscillographic display, boundaries were demarcated by placing cursors at zero crossings at the onset and offset of periodicity corresponding to the vowel, and the duration (in milliseconds) between the cursors was computed.

2) *Keyword F_0* - to characterize changes in F_0 throughout the utterance, the fundamental frequency of the stressed vowel of each content word was computed. In addition, the terminal portion of the F_0 contour (believed critical in signaling distinctions in linguistic modality) was calculated within the final 150 msec of each utterance. By means of visual inspection of the waveform, five contiguous pulses were isolated at the centre of the vowel or within the final 150 msec of the utterance by placing cursors at zero crossings. The inverted period of the five pulses was then averaged to derive the mean F_0 (Behrens, 1988; Ouellette & Baum, 1993).

3) *Focus Accent* - the extent of local fundamental frequency change associated with sentence focus was determined at the two locations where items could receive emphasis in Experiment 1: sentence-initially and sentence-finally. As focused words were always bisyllabic, focus accent was calculated as the difference between the mean F_0 of the initial stressed vowel (determined in (2) above) and the mean F_0 of the terminal portion of the second vowel (e.g., *Ma-ry*).

"Global" Measures

1) *Speech Rate* - as discrete emotions may be distinguished by differences in speaking rate, speech rate was calculated as the number of syllables present divided by the total sentence duration.

2) *Mean F_0* - the mean F_0 of each utterance was calculated as the average of *all* F_0 values derived at keyword locations within the utterance including the utterance terminal. Thus, four values contributed to the utterance mean for short utterances, whereas five values contributed to the mean for long utterances.

3) *F_0 Range* - as F_0 range may be a particularly important feature in the expression of emotion, the F_0 range of each utterance was computed as the difference between the highest and the lowest F_0 values derived at all keyword locations within the utterance including the utterance terminal.

Statistical Analyses

Acoustic measures were normalized prior to statistical analysis. To adjust for inter-speaker differences in speaking rate, keyword duration values were divided by the corresponding utterance duration for each speaker. To normalize for differences in F_0 range between subjects, F_0 values (keyword F_0, utterance terminal, utterance mean F_0) were transformed to *z* scores within speakers using the following formula: $F_{0norm} = (F_{0i} - F_{0mean})/ s$, where F_{0i} is the observed F_0 value, F_{0mean} is the mean F_0 across all utterances produced by the speaker (both short and long), and *s* is the standard deviation (Colsher et al., 1987; Gandour et al., 1995). Measures of F_0 range (focus accent, utterance range) were normalized for inter-subject differences in mean F_0 by dividing the range values by the mean F_0 of the corresponding utterance for each speaker.

RESULTS

The normalized data from Experiment 1 were examined statistically by means of analysis of variance (ANOVA) techniques, independently for each utterance length (short, long). Data derived for the two "short" items and the two "long" items were collapsed within each condition prior to statistical inspection. In reporting statistically significant effects, findings representative of the short and long stimuli are presented concurrently, followed by exceptional patterns of performance, where observed. Statistically significant effects were explored post hoc using Tukey's HSD procedure ($p<.05$), wherever appropriate.

1) Local "Keyword" Measures

Following Cooper, Eady and their associates (Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986), measures of keyword duration and keyword F_0 were computed to arrive at a comprehensive profile of acoustic change occurring throughout the utterances elicited in Experiment 1. In addition, a measure of the magnitude of F_0 change associated with focused items ("F_0 accent") was determined at assigned keyword locations. By these means, the RHD patients' proficiency at modulating prosody locally in sentences differing in affective mode and overall length may be illuminated. As well, these data may prove insightful of the acoustic interaction between highly localized and more gradual changes in prosodic content in normal speech production.

Data obtained for the keyword duration and keyword F_0 analyses were explored using mixed ANOVAs in a full factorial design, performed separately on utterances varying in length (short, long) and affect type (neutral, sad, happy, angry) for a total of 8 distinct analyses per measure. For keyword duration, four 2 X 2 X 3 X 3 and four 2 X 2 X 3 X 4 analyses were conducted for the short and long stimuli, respectively. For keyword F_0, four 2 X 2 X 3 X 4 and four 2 X 2 X 3 X 5 ANOVAs were performed for the short and long stimuli, respectively. For all 16 analyses, GROUP (NC, RHD) served as the between-subjects factor, and MODALITY (declarative, interrogative), FOCUS (none, initial, final), and KEYWORD position (defined below for each acoustic measure) served as the repeated factors. Only the number of levels assigned to the KEYWORD factor distinguished the analyses conducted on the keyword duration and keyword F_0 data, depending on the length of the stimulus and whether "terminal" measures were extracted (these were computed only for keyword F_0).

The results of keyword analyses are reported together for both sentence lengths but individually according to affect type, to underscore potential differences in each group's ability to modulate local linguistic parameters as a function of affective speaking mode. However, as many patterns in the stimuli displayed properties independent of affective mode or sentence length for both groups, effects characteristic of *all* stimuli are described in an initial section within the discussion of each keyword measure.

(i) Keyword Duration

Temporal Features Characteristic of All Stimuli

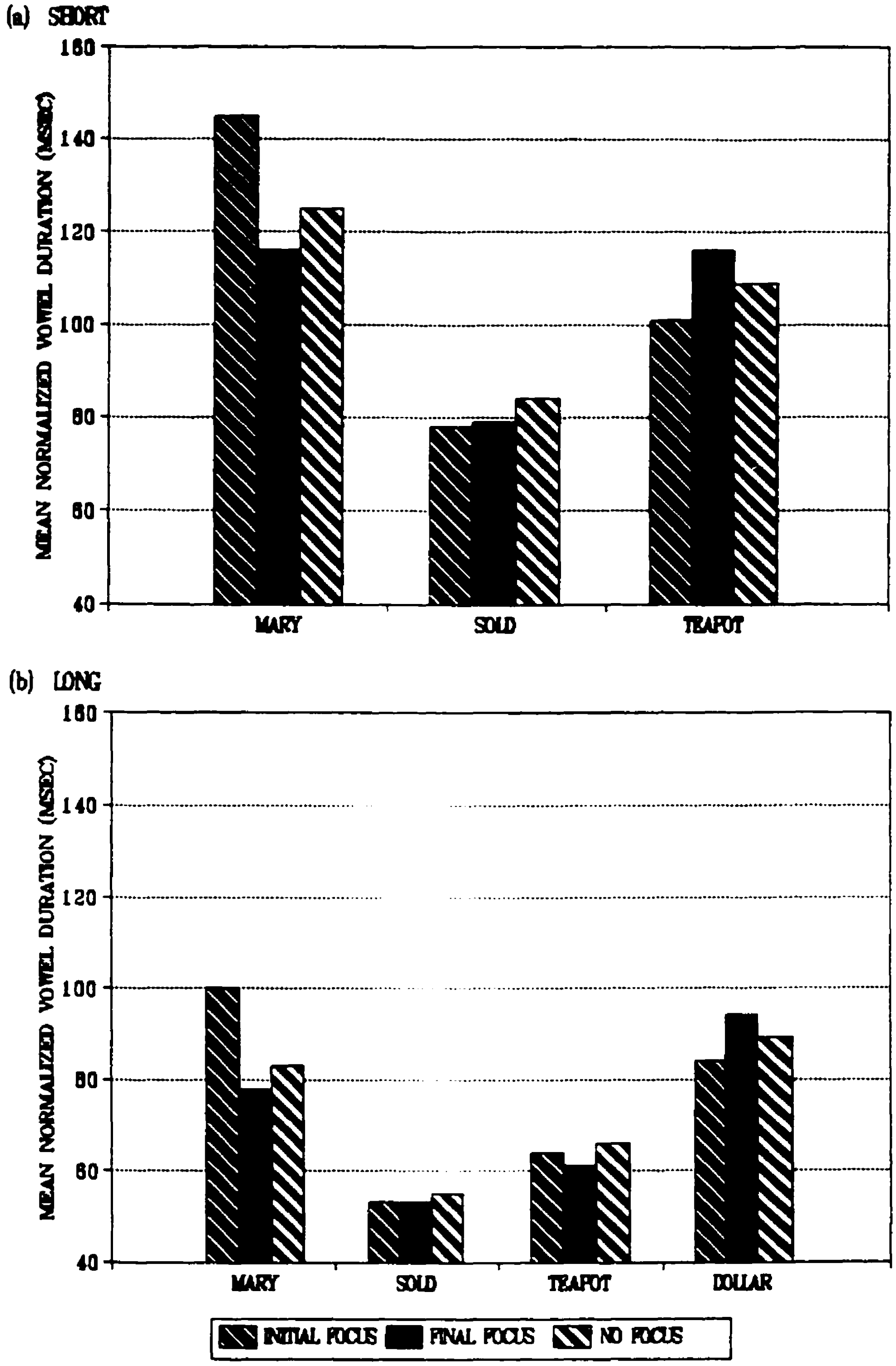
Keywords for which vowel durations were determined in short (three keywords) and long (four keywords) utterances are reviewed in Table 2. The outcome of the eight four-way (GROUP X MODALITY X FOCUS X KEYWORD) ANOVAs performed on the keyword duration data pointed to temporal distinctions common to all stimuli, irrespective of emotion type or sentence length. Namely, all eight analyses produced significant main effects for FOCUS[2] and KEYWORD, and a significant interactive effect of FOCUS X KEYWORD (p<.05 in all cases--see Appendix A for individual F values). A graphic display of the two-way interaction is presented in Figure 1 (a-b), independently by sentence length (collapsed across the four emotions).

Table 2. Position of keywords in short and long utterances where vowel durations were computed.

"Short"		K1	K2	K3	
	e.g.,	*Mary*	*sold*	*the teapot*	
"Long"		K1	K2	K3	K4
	e.g.,	*Mary*	*sold*	*the teapot*	*for a dollar*

[2]A main effect for FOCUS was marginally significant for short stimuli when angry prosody was elicited (p=.055).

Figure 1. Interaction of focus and keyword position for (a) short and (b) long stimuli, collapsed across the four emotions.



Post hoc inspection of these interactions initially considered how differences in focus location influenced vowel durations at each keyword position. For sentence-initial position (K1), focused vowels were always significantly longer than unfocused vowels (i.e., utterances with either no focus or sentence-final focus). For the neutral stimuli (and short, sad stimuli), sentence-initial vowels were also longer in tokens with no focus than those with focus located sentence-terminally, indicating that focus produced toward the end of these stimuli was anticipated through reduction in initial vowel duration.

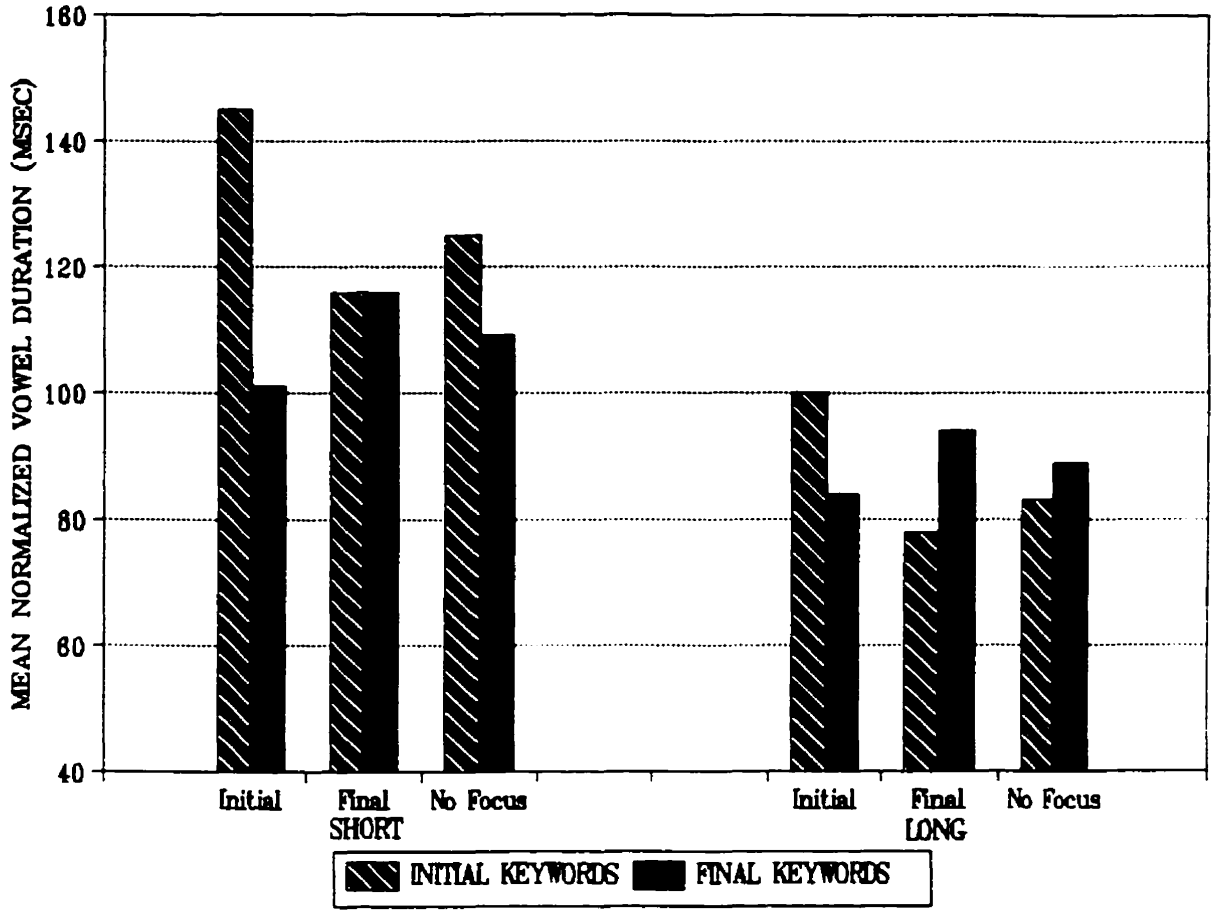
In sentence-final position (K3 for short, K4 for long), focused items were significantly longer than corresponding items in utterances with sentence-initial focus for all eight analyses. Focused vowels in sentence-final position were also significantly greater in duration than corresponding vowels in utterances without focus, although again, this was only true for the neutral stimuli (short and long). This pattern suggests that, despite similar overall patterns, focus expressed in a neutral tone may have resulted in greater temporal distinctions in both sentence-initial and sentence-final position than focus expressed in various affective tones. Vowel durations at "intervening" keyword positions (i.e., those between sentence-initial and sentence-final position) were invariably unaffected by differences in focus placement.

Further exploration of these interactions looked at differences in keyword durations within each level of focus. For all focus versions, adjacent keywords always differed significantly in duration for both short and long utterances: K1 and K3 always exceeded K2 (which displayed the shortest durations overall); for the long stimuli, K3 was always shorter than K4. However, as expected, comparison of the two keyword positions in which focus was typically expressed (sentence-initial and sentence-final) revealed different patterns of duration as a function of focus location. This relationship is expressed in Figure 2 for both the short and long stimuli.

For the short stimuli, sentence-initial vowels were significantly longer than sentence-final vowels for utterances with initial or no focus. However, durations in these two positions did not significantly differ when focus was produced sentence-finally, suggesting that the *absence* of vowel elongation at the beginning of utterances may have been important in signaling final focus for the short stimuli. For the long stimuli, a somewhat different pattern emerged; the presence of focus in either sentence-initial or sentence-final position always led to a significant duration increase in that position relative to the unfocused position. However, when utterances lacked focus, no consistent relationship was observed between these two keyword durations (K1 did not differ from K4 for neutral and sad prosody, and K4 exceeded K1 for happy and angry prosody). These data suggest that local changes in duration marked directly on focused items occurred more frequently as utterances increased in length (review Figure 2 and see further discussion below).

Other significant effects proved common only to the four analyses performed on either the short or the long data. For short utterances but not long

Figure 2. Comparison of sentence-initial and sentence-final keyword durations for the short and long stimuli.

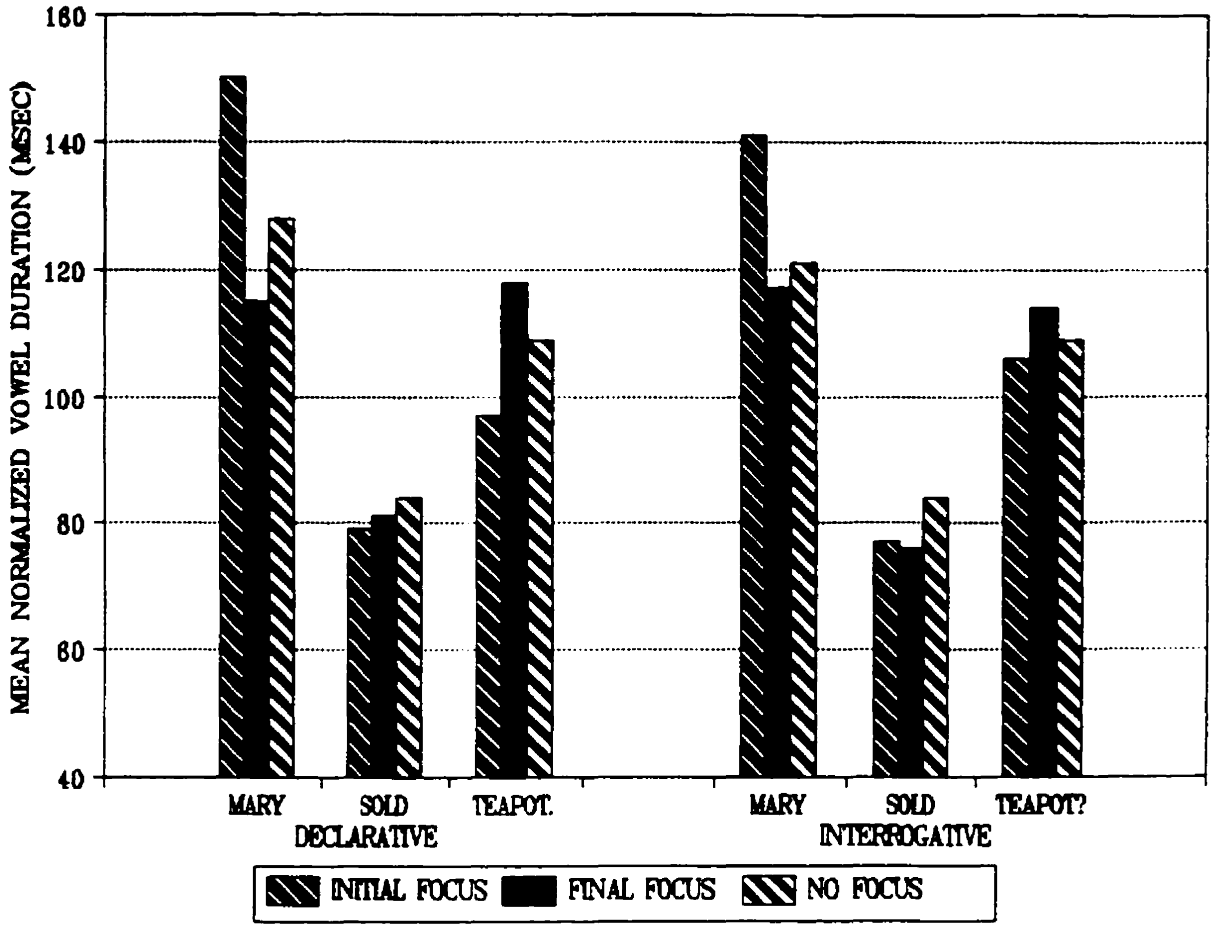


utterances, a significant interactive effect of MODALITY X FOCUS X KEYWORD was always observed ($p<.05$ in all cases--see Appendix A). This finding suggests that the interaction of focus location and keyword duration (described above for both short and long stimuli) was further dependent on the linguistic modality of utterances when sentences were relatively short. The three-way interaction is depicted graphically in Figure 3.

Investigation of the interaction assessed how differences in sentence modality influenced the relationship between keyword duration and focus location. For declarative intonation, focused items were always significantly longer than unfocused items (with sentence-final or no focus) when produced sentence-initially. Moreover, sentence-initial vowels displayed *reduced* durations in utterances with final emphasis relative to those without emphasis for three of the four affect types (neutral, sad, angry). On sentence-final keywords in declaratives, vowels were always significantly longer when emphasized than when emphasis was located in sentence-initial position. In general, the pattern described for short declarative utterances reflects the overall relationship between focus and keyword duration described for both short and long stimuli above.

Analysis of short utterances spoken with interrogative intonation yielded a somewhat different outcome. Initial focused items were again signficantly longer than unfocused items (with sentence-final or no focus) when produced at the beginning of the sentence. However, differences in vowel length did not serve to distinguish items with or without focus in sentence-final position when

Figure 3. Interaction of sentence modality, focus, and keyword position for the short stimuli.



interrogative intonation was produced. These data suggest that the presence of interrogative intonation may have constrained the subjects' ability to modulate temporal properties of sentence-final items when relatively short utterances were produced (review Figure 3).

Neutral (Non-affective) Stimuli

The subjects' ability to produce neutral prosody was of considerable theoretical interest in the present study: these data served to demonstrate the RHD patients' capacity to modulate prosodic patterns in a non-affective context, and permitted a re-examination of how differences in sentence modality and focus location influence the acoustic form of neutral utterances for normal subjects (Cooper et al., 1985; Eady & Cooper, 1986; Ryalls et al., 1994). Table 3 supplies the mean normalized keyword durations computed for the neutral stimuli for each group, as a function of sentence modality and focus location (by sentence length).

As may be seen, vowel durations tended to be greater in short utterances than in long utterances for both groups irrespective of stimulus type. In addition to significant effects common to all stimuli (described above), separate four-way (GROUP X MODALITY X FOCUS X KEYWORD) ANOVAs performed on the short and long data yielded a significant interaction of GROUP X MODALITY X KEYWORD for the neutral stimuli [$F_{SHORT}(2,36) = 4.22, p<.05$; $F_{LONG}(3,54) = 6.58, p=.001$]. This interaction is illustrated in Figure 4, independently for the short and long utterances.

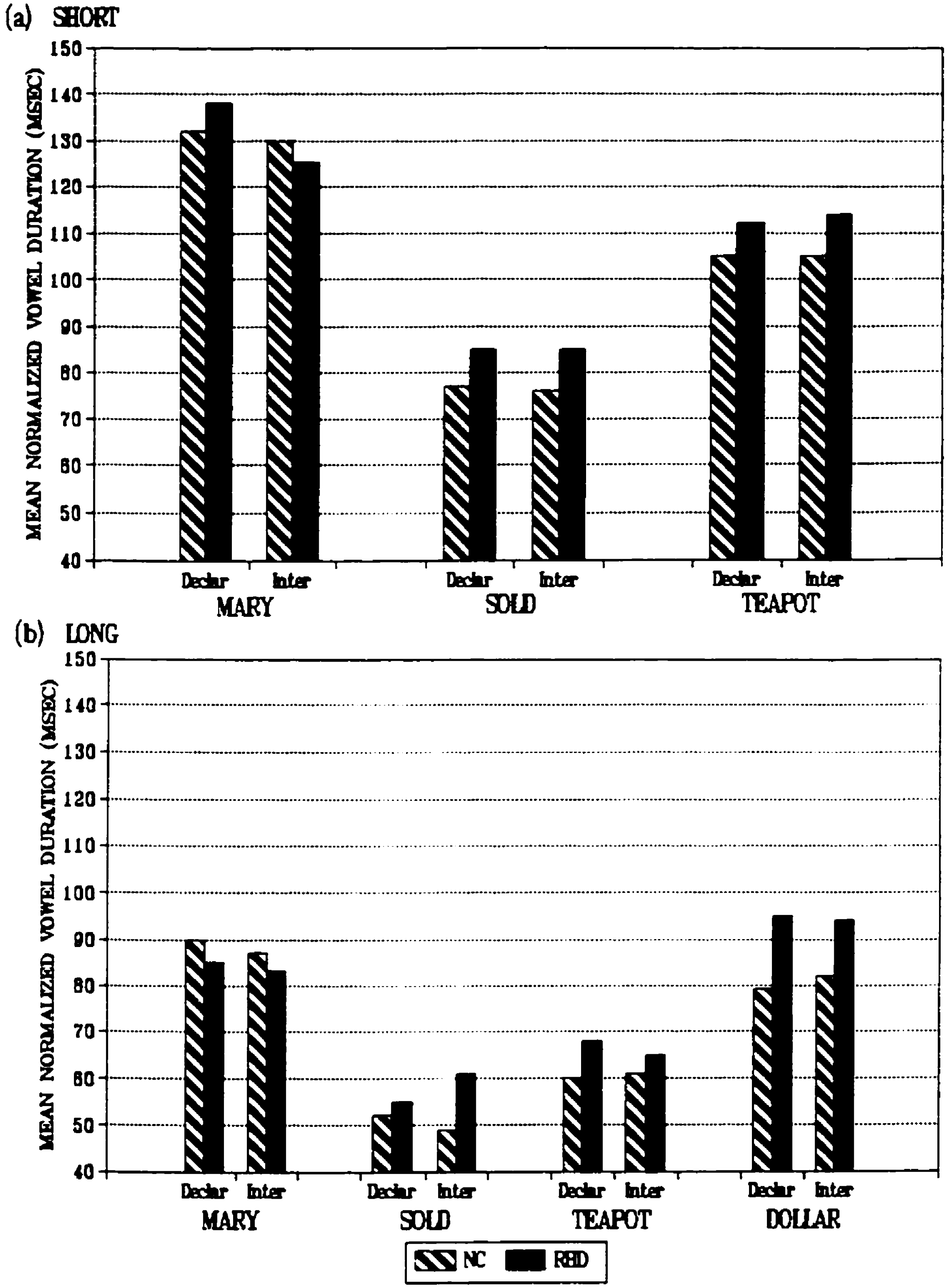
Table 3. Neutral Stimuli - Mean normalized keyword durations for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

			K1 MARY		K2 SOLD		K3 TEAPOT		K4 DOLLAR	
Length	*Group*	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Short	NC	Initial	151	157	76	74	94	101	--	--
		Final	113	120	76	74	116	109	--	--
		None	133	114	80	79	106	104	--	--
	RHD	Initial	155	130	83	85	100	112	--	--
		Final	120	119	87	81	128	120	--	--
		None	138	127	86	88	107	111	--	--
Long	NC	Initial	108	104	48	49	57	62	69	79
		Final	77	77	51	51	60	61	88	85
		None	86	81	56	48	62	60	80	81
	RHD	Initial	96	95	56	61	69	66	91	91
		Final	74	73	55	59	63	62	103	98
		None	84	82	55	64	72	66	91	92

Note: (.)=declarative, (?)=interrogative. Values were normalized for intra-speaker differences in speaking rate and then multiplied by 1000 for ease of presentation.

Comparison of each group's performance as a function of keyword position and sentence type indicated that for the short tokens, RHD patients produced significantly longer vowel durations than normal subjects for all keyword items in both modalities, except for sentence-initial items in interrogatives (which did not

Figure 4. Interaction of group, sentence modality, and keyword position for the (a) short and (b) long stimuli spoken in a neutral context.



significantly differ between groups). Fewer group differences were noted for the long tokens, although RHD subjects produced significantly longer durations than normal subjects on sentence-final keywords (both declaratives and interrogatives) and on items at two intervening positions (K3 for declaratives, K2 for interrogatives). Keyword durations produced by the NC subjects did not exceed those of the RHD patients for any comparison for either stimulus length.

Comparison of keyword durations at sentence-initial versus sentence-final position within each group indicated that sentence-initial items were always significantly longer than sentence-final items for the normal subjects, and for the RHD subjects when producing short sentences. For long sentences, however, this pattern was reversed for the RHD subjects, with vowels in sentence-final position proving significantly longer than those in sentence-initial position. Items in intervening positions always displayed significantly reduced vowel durations relative to items in sentence-initial and sentence-final position for both groups. Generally, differences in sentence modality were not reflected in significant vowel duration differences at any keyword position for either group.[3]

Significant effects related specifically to the production of short or long utterances also emerged. For short utterances, interactions of MODALITY X KEYWORD [$F(2,36)=6.84, p<.01$] and MODALITY X FOCUS X KEYWORD (reported in Appendix A) were observed; as described earlier, the three-way

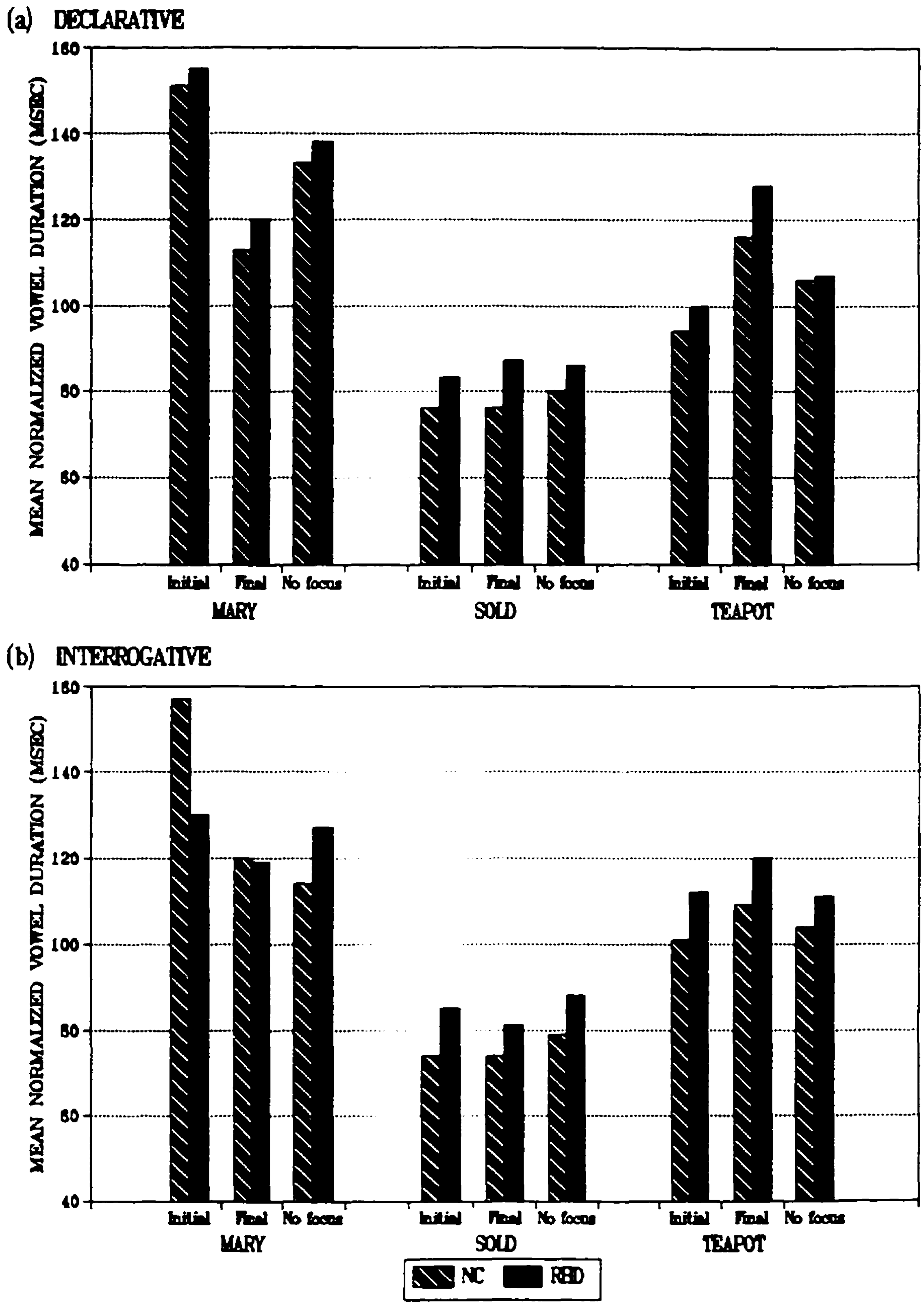
[3]Exceptionally, the RHD patients produced significantly longer vowels in sentence-initial position for declarative as opposed to interrogative utterances, only for the short stimuli.

interaction constituted a temporal effect common to all short stimuli. As well, interactions of GROUP X FOCUS [$F(2,36)=5.2$, $p=.01$], GROUP X FOCUS X KEYWORD [$F(4,72)=2.80$, $p<.05$], and a four-way interaction of GROUP X MODALITY X FOCUS X KEYWORD [$F(4,72)=3.02$, $p<.05$] emerged from the analysis of the short stimuli. Figure 5 (a-b) depicts the four-way interaction, independently by sentence modality.

Follow-up tests performed on the interaction indicated that for each group, focused items were significantly longer than unfocused items in both sentence-initial and sentence-final position when declarative intonation was produced. However, for interrogative intonation (where focus is generally less distinctive in short utterances based on vowel duration--see above), the NC subjects used duration to distinguish focused and unfocused items only in sentence-initial position. Right-brain-damaged patients failed altogether to produce significant temporal distinctions to focus in either sentence-initial or sentence-final position when producing interrogative intonation.

Finally, when long utterances were examined, significant effects of GROUP [$F(1,18)=6.63$, $p<.05$] and GROUP X KEYWORD [$F(3,54)=4.60$, $p<.01$] were produced. As described earlier (and as shown in Figure 2), this interaction represented the RHD patients' tendency to produce longer vowel durations in sentence-final position than the NC subjects, overall.

Figure 5. Interaction of group, sentence modality, focus, and keyword position for short (a) declaratives and (b) interrogatives spoken in a neutral context.



Sad Stimuli

Eliciting the experimental stimuli from subjects in various affective modes was necessary to explore how emotional and non-emotional prosody are produced in tandem by NC and RHD subjects; this design represents a departure from previous studies in which keyword values were determined only for non-affective stimuli (Cooper et al., 1985; Eady & Cooper, 1986; Ryalls et al., 1994). Keyword durations computed for the NC and RHD subjects when utterances were spoken in a sad context are provided in Table 4, as a function of sentence modality and focus location (by sentence length).

In addition to "common" temporal patterns in the keyword data, analysis of the sad stimuli produced a main effect for GROUP [$F(1,18)=10.00, p<.01$] and a significant GROUP X FOCUS X KEYWORD interaction [$F(6,108)=2.42, p<.05$] for the long utterances. Closer inspection of the interaction, displayed in Figure 6, indicated that the RHD patients signaled focus in sentence-initial position with significantly shorter vowels than the NC subjects. Moreover, keywords immediately following sentence-initial position (i.e., K2) were significantly *elongated* by the RHD patients relative to the NC subjects when producing long sentences, rendering initial focused items far less distinctive in the RHD subjects' speech relative to that of the NC subjects.[4]

---

[4]The RHD patients also produced significantly greater durations than the NC subjects on penultimate keywords (K3) for utterances with final focus, and on sentence-final keywords (K4) for utterances without focus.

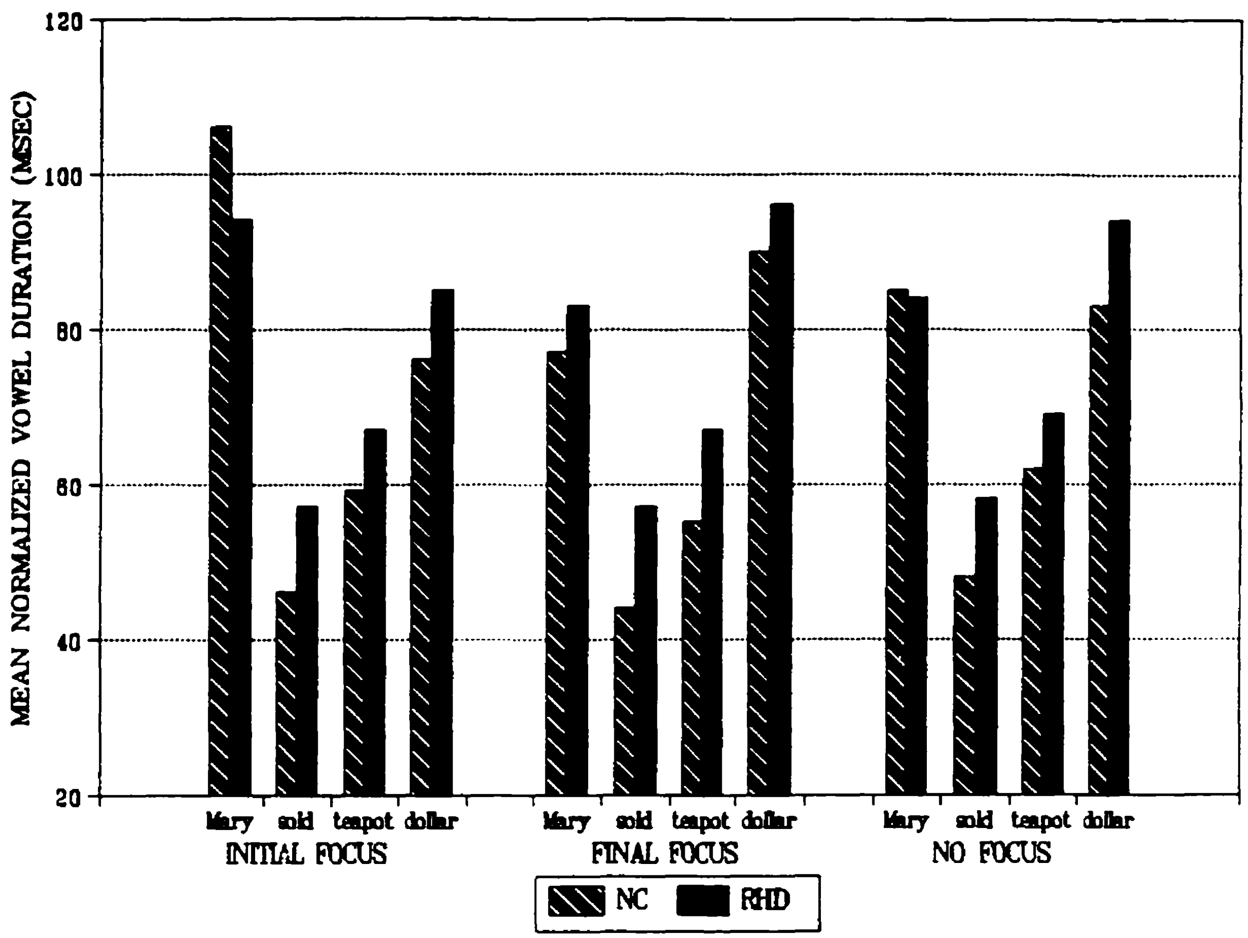
Table 4. Sad Stimuli - Mean normalized keyword durations for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

			K1		K2		K3		K4	
			MARY		SOLD		TEAPOT		DOLLAR	
Length	*Group*	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Short	NC	Initial	152	154	76	73	89	96	--	--
		Final	115	117	79	68	111	101	--	--
		None	125	123	82	79	97	102	--	--
	RHD	Initial	149	133	77	81	99	112	--	--
		Final	114	118	88	82	115	118	--	--
		None	130	129	91	92	109	108	--	--
Long	NC	Initial	104	108	48	43	59	59	74	78
		Final	78	75	46	42	55	55	88	91
		None	86	84	49	47	62	61	81	84
	RHD	Initial	86	91	58	56	65	68	88	82
		Final	86	80	57	57	66	67	96	95
		None	83	85	56	60	69	69	94	94

Note: (.)=declarative, (?)=interrogative. Values were normalized for intra-speaker differences in speaking rate and then multiplied by 1000 for ease of presentation.

Comparison of sentence-initial and sentence-final keyword durations for the sad stimuli indicated that the NC subjects always lengthened vowels in the position receiving focus relative to the alternate position, consistent with the general pattern described for all long stimuli. In contrast, the RHD patients produced this distinction only for items receiving focus sentence-terminally, thus

Figure 6. Interaction of group, focus, and keyword position for long stimuli spoken in a sad context.



failing to provide the normal cues to sentence-initial focus. It should be noted that no group interactions were found in the analysis of sad stimuli when *short* utterances were examined.

Happy Stimuli

Keyword durations calculated for the short and long utterances elicited in a happy context are presented in Table 5, as a function of linguistic modality and focus distribution. In addition to common effects (reported in Appendix A), significant effects of GROUP [$F(1,18)=8.14$, $p<.05$] and GROUP X MODALITY X FOCUS [$F(2,36)=4.40$, $p<.05$] were observed when long sentences were elicited in a happy tone. Comparison of group performance as a function of modality and focus location indicated that RHD patients produced significantly greater vowel durations than normal subjects, overall, for four of the six modality/focus combinations (declaratives with initial and final focus, interrogatives with final and no focus). For only the RHD patients, keyword durations averaged across the utterance tended to be greater in declaratives with initial focus than in other declarative versions, or in interrogatives with final focus. For only the NC group, long interrogatives with initial focus exhibited significantly greater keyword durations than interrogatives with final focus, overall. Keyword durations in utterances matched for focus location did not significantly differ as a function of sentence modality for either group. Again, no group interactions emerged from the analysis of happy stimuli when *short* utterances were elicited.

Table 5. Happy Stimuli - Mean normalized keyword durations for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

			K1		K2		K3		K4	
			MARY		SOLD		TEAPOT		DOLLAR	
Length	*Group*	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Short	NC	Initial	151	142	81	68	97	98	--	--
		Final	116	113	72	69	119	115	--	--
		None	121	124	80	80	113	109	--	--
	RHD	Initial	142	134	90	83	107	108	--	--
		Final	118	116	84	85	117	115	--	--
		None	122	122	87	89	117	114	--	--
Long	NC	Initial	103	101	49	49	59	63	78	84
		Final	75	75	49	45	60	57	92	90
		None	81	81	55	48	66	65	88	86
	RHD	Initial	94	96	64	58	74	67	95	94
		Final	77	79	62	59	67	66	98	99
		None	80	84	62	61	66	72	94	97

Note: (.)=declarative, (?)=interrogative. Values were normalized for intra-speaker differences in speaking rate and then multiplied by 1000 for ease of presentation.

Angry Stimuli

Keyword durations computed for the short and long stimuli elicited in an angry context are supplied for the two subject groups in Table 6, by sentence modality and focus location. The (GROUP X MODALITY X FOCUS X

KEYWORD) ANOVAs performed on the short and long utterances elicited in an angry tone yielded interactions of MODALITY X KEYWORD [$F(2,36)=7.14$, $p<.01$] and GROUP X FOCUS X KEYWORD [$F(4,72)=2.98$, $p<.05$] for the short stimuli. The interaction of modality and keyword was discussed within the consideration of the three-way (MODALITY X FOCUS X KEYWORD) interaction described for all stimuli (reported in Appendix A). The interaction of group, focus, and keyword is depicted in Figure 7.

Comparison of group performance as a function of keyword and focus location established that the RHD patients produced significantly reduced vowel durations relative to the NC subjects on focused sentence-initial items. No further between-groups comparisons proved significant. The ability of each group to employ duration to signal focus location at designated keyword positions revealed further differences of note; although the NC group produced temporal distinctions conforming to the pattern described for short stimuli in general (i.e., initial focused items were significantly longer than all corresponding unfocused items, and final focused items were significantly longer than corresponding items in utterances with initial focus), the RHD group produced relatively few distinctions. Only focused items in initial position differed significantly from sentence-initial items in utterances with final focus (see Figure 7).

Table 6. Angry Stimuli - Mean normalized keyword durations for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

			K1		K2		K3		K4	
			MARY		SOLD		TEAPOT		DOLLAR	
Length	*Group*	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Short	NC	Initial	155	148	69	70	90	102	--	--
		Final	112	117	73	68	117	113	--	--
		None	119	113	73	79	108	107	--	--
	RHD	Initial	141	128	79	80	98	113	--	--
		Final	112	109	83	77	125	114	--	--
		None	129	117	89	85	114	111	--	--
Long	NC	Initial	103	109	45	44	57	60	77	77
		Final	77	77	44	44	54	57	90	84
		None	84	81	46	44	59	61	85	88
	RHD	Initial	95	95	59	54	68	66	91	89
		Final	81	79	58	61	62	63	101	101
		None	83	83	60	59	73	68	95	93

Note: (.)=declarative, (?)=interrogative. Values were normalized for intra-speaker differences in speaking rate and then multiplied by 1000 for ease of presentation.

Effects arising solely from the analysis of long utterances included a significant main effect for GROUP [$F(1,18)=16.50$, $p=.001$] and a significant GROUP X KEYWORD interaction [$F(3,54)=3.34$, $p<.05$]. Inspection of the interaction (presented in Figure 8) revealed that the RHD subjects failed to

Figure 7. Interaction of group, focus, and keyword position for short stimuli spoken in an angry context.

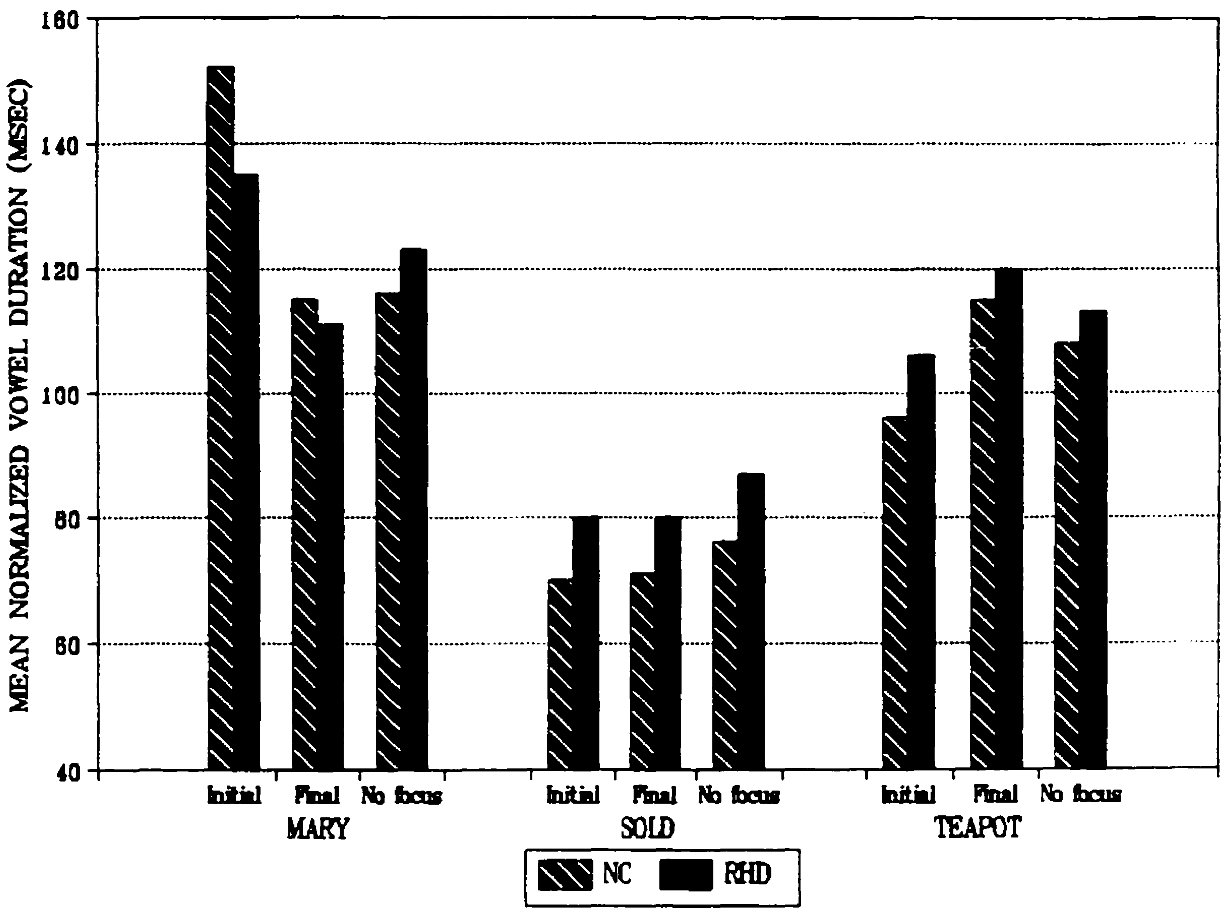
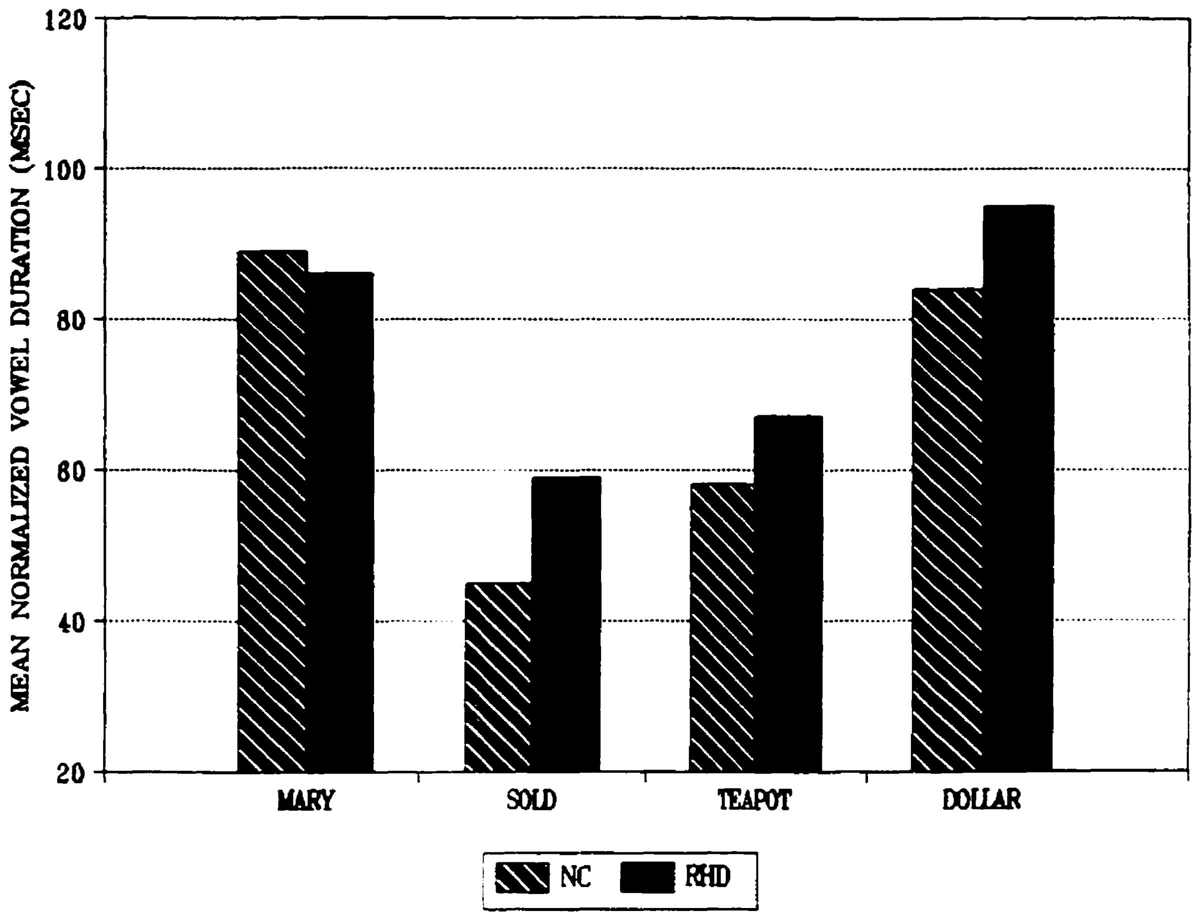


Figure 8. Interaction of group and keyword position for long stimuli spoken in an angry context.



reduce vowels occurring immediately following sentence-initial position (i.e., K2) as much as did the normal subjects. This resulted in significantly greater durations in this position for the RHD patients relative to the NC subjects. As well, in contrast to the NC subjects, there were no significant duration differences between intervening keyword items (i.e., K2 and K3) for the RHD group.

Summary of Keyword Duration Data

To summarize the major findings of the keyword duration data, irregularities were noted in the production of prosodic stimuli by RHD patients relative to non-neurological control subjects. In general, these abnormalities could be characterized by the tendency of RHD patients to provide fewer (or smaller) duration cues to emphasis contrasts than NC subjects. This pattern was particularly (but not exclusively) true for long utterances, for which group effects emerged for all four affect types (group interactions were evident only in the neutral and angry contexts for short stimuli). Temporal irregularities were not confined to the production of affective speech by the RHD adults, as difficulties were consistently prevalent in the neutral context as well. More generally, it was shown that short and long utterances place somewhat different demands on the normal speaker *vis-à-vis* the temporal marking of linguistic focus, short utterances involving temporal dependencies not necessarily localized to the focused item.

(ii) Keyword F_0

A point-by-point analysis of mean F_0 at designated keyword positions was conducted to further explore each group's capacity to produce prosodic distinctions in various contexts. In addition to the keywords identified for duration measures, the terminal F_0 of all utterances was additionally computed (measured within the final 150 msec of the utterance). Table 7 summarizes the location at which mean F_0 measures were extracted in the short and long stimuli.

The eight (GROUP X MODALITY X FOCUS X KEYWORD) ANOVAs, performed separately for each affect type and for the short and long stimuli, yielded a number of common effects. As before, F_0 patterns characteristic of all stimuli are described initially, followed by patterns in the data that were uniquely related to the production of prosody in specific affective modes.

Table 7. Position of keywords in short and long utterances where mean F_0 was computed.

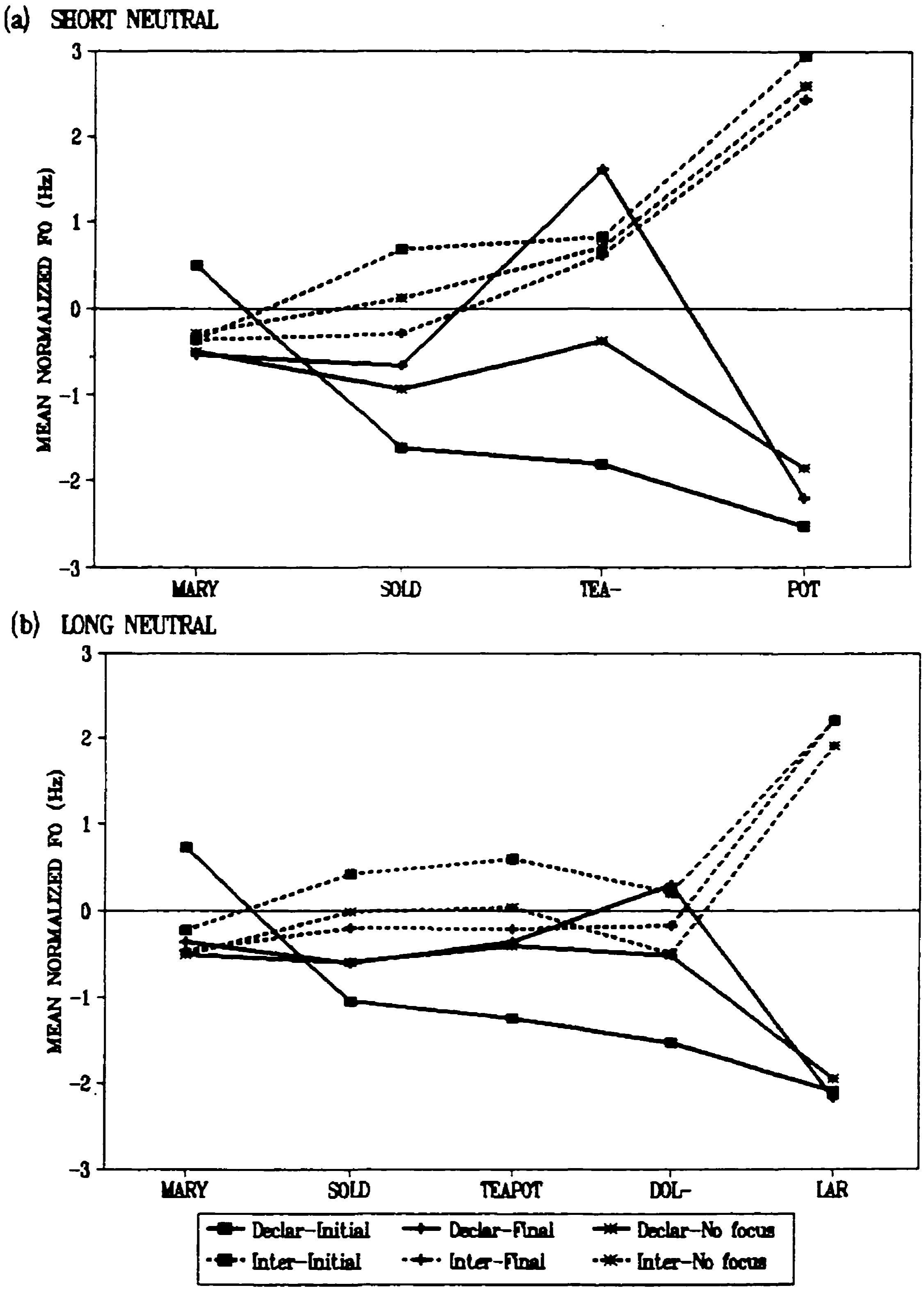
"Short"		K1 K2 K3 K4
	e.g.,	*Mary* *sold* *the* ***tea-pot***
"Long"		K1 K2 K3 K4 K5
	e.g.,	*Mary* ***sold*** *the* *teapot for a* ***dol-lar***

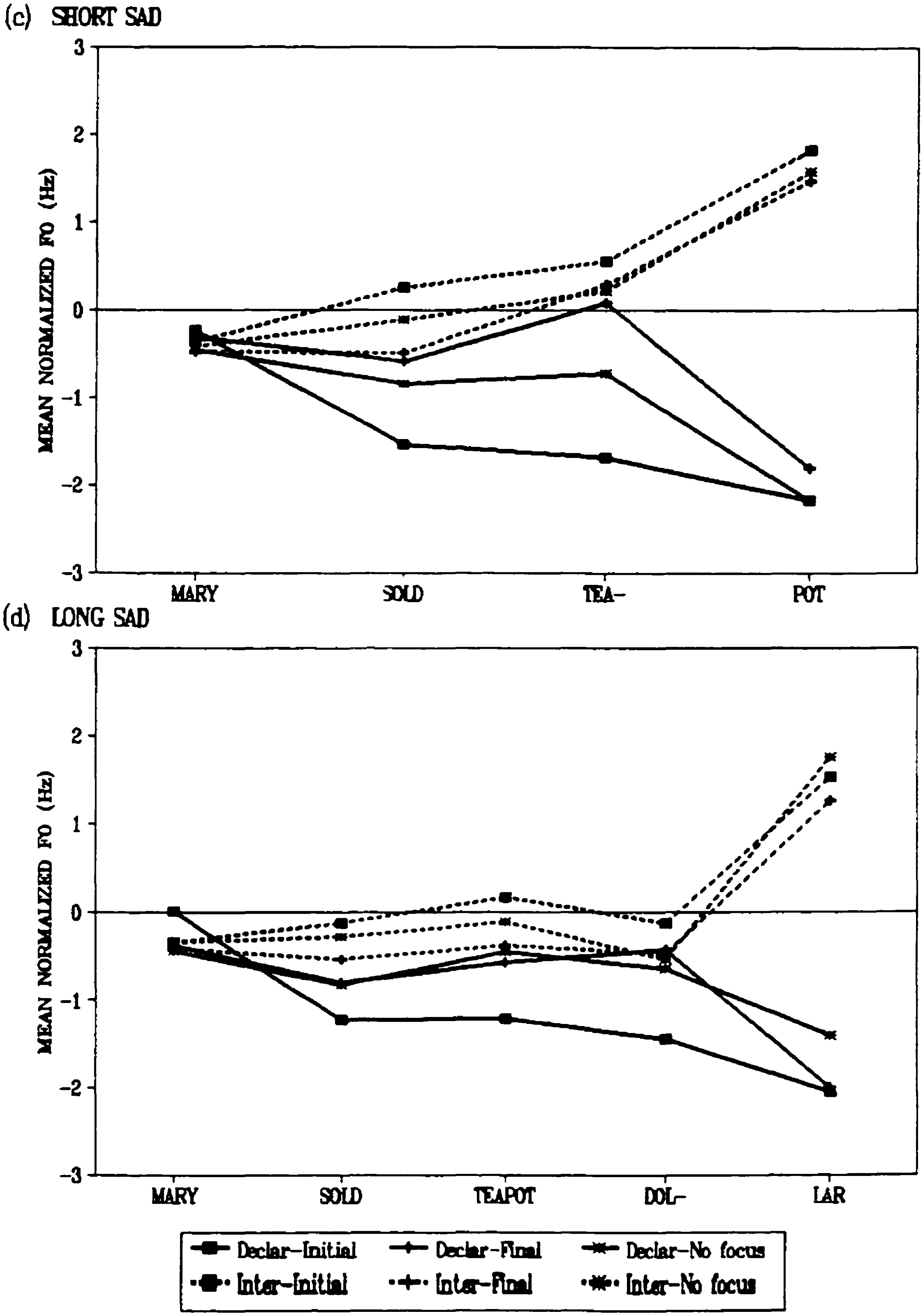
F_0 Features Characteristic of All Stimuli

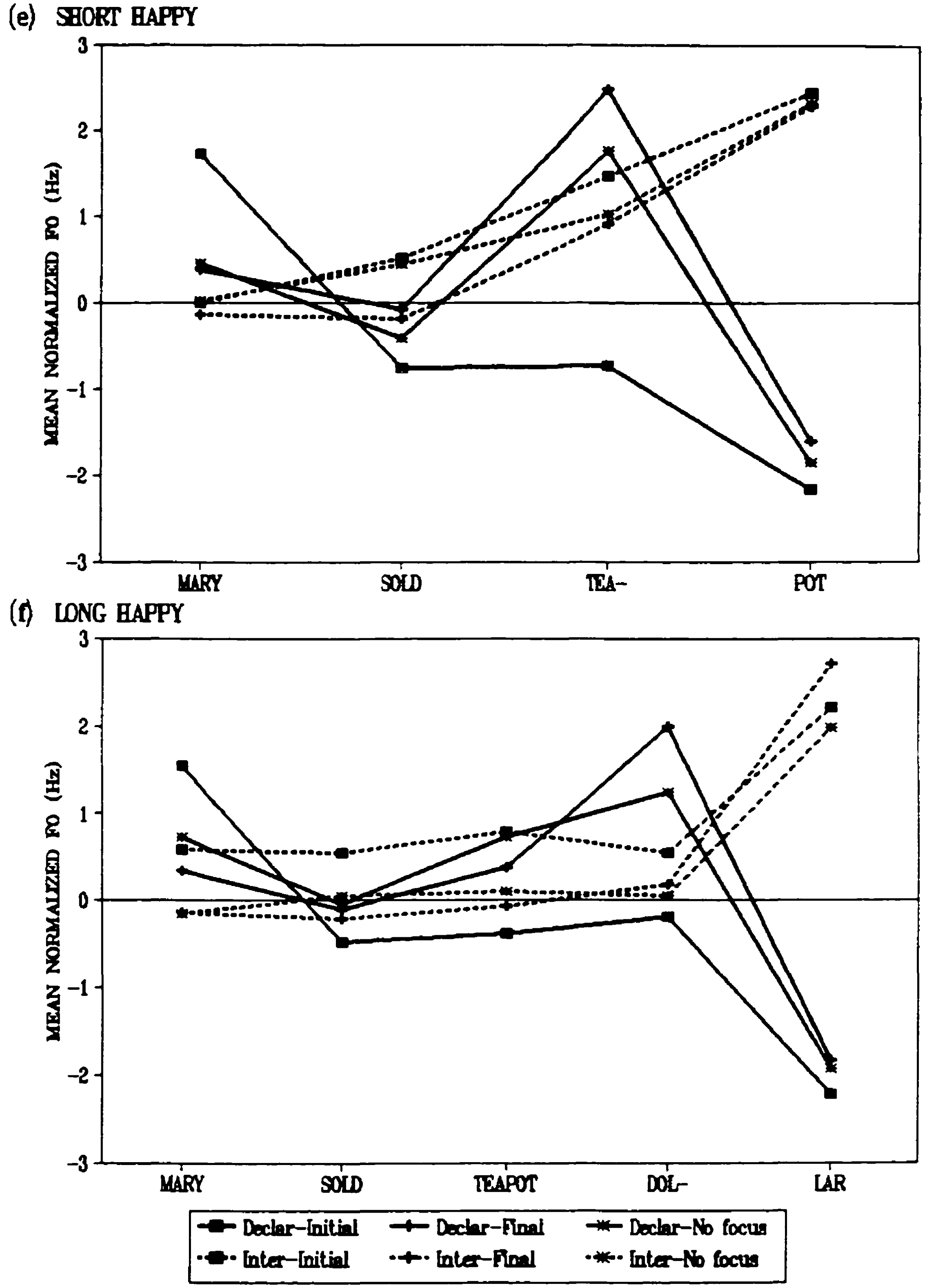
The eight analyses performed on the keyword F_0 data each yielded the following significant effects: MODALITY, MODALITY X KEYWORD, FOCUS X KEYWORD, MODALITY X FOCUS, and MODALITY X FOCUS X KEYWORD ($p<.05$ in all cases--*F* values are listed in Appendix x). The three-way interaction is displayed graphically in Figure 9 (a-h) for the neutral (a-b), sad (c-d), happy (e-f), and angry (g-h) stimuli, independently by sentence length.

Inspection of the three-way interaction first examined differences in mean F_0 at adjacent keyword positions for specific combinations of focus and sentence modality. For both short and long utterances, focus in sentence-initial position led to a significant decrement in F_0 on all subsequent items when declarative intonation was produced. When interrogatives were produced, however, focus in sentence-initial position was marked for only one stimulus type (short utterances with neutral prosody), in which case a significant *rise* was noted on all subsequent items (Eady & Cooper, 1986; see Figure 9a). For utterances with sentence-final emphasis, a significant rise in F_0 was observed on the focused item (K3 for short, K4 for long) only in the case of *short, declarative utterances* (all four emotions). In long utterances, sentence-final items were not marked by a significant rise in F_0, while interrogative utterances (short and long) were marked in sentence-final position for only one of the eight stimulus types (again, short neutral stimuli displayed a significant rise in F_0 on the focused item--see Figure 9a).

Figure 9. Interaction of sentence modality, focus, and keyword position for short and long stimuli spoken in a neutral (a-b), sad (c-d), happy (e-f), and angry (g-h) context.

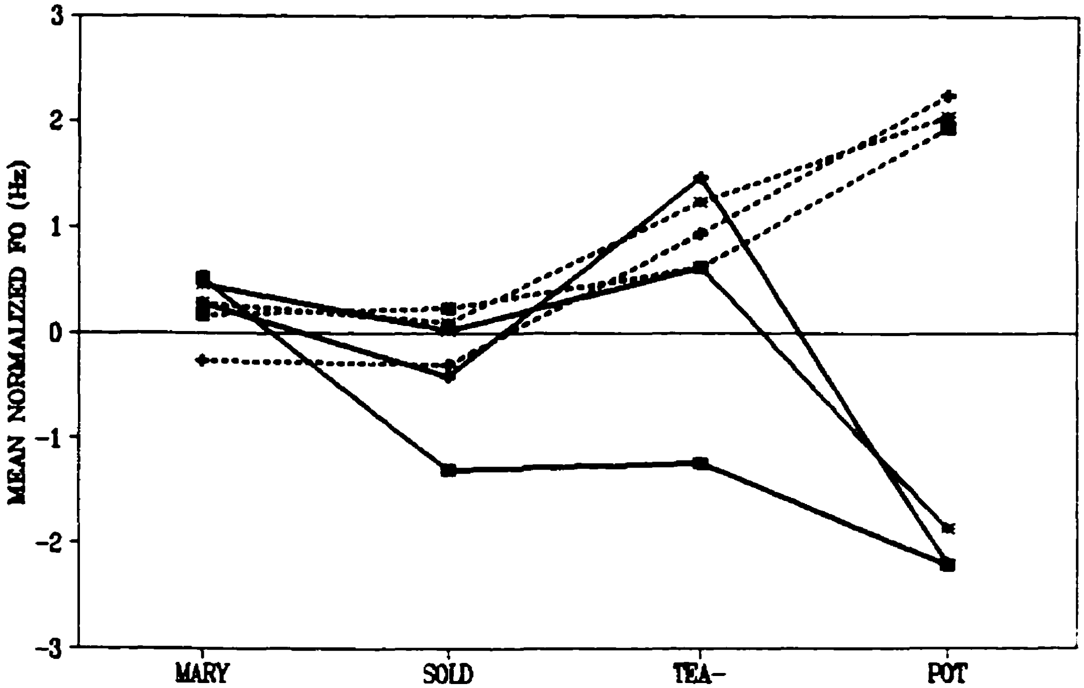



(c) SHORT SAD
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEA-
POT
(d) LONG SAD
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEAPOT
DOL-
LAR
Declar-Initial
Declar-Final
Declar-No focus
Inter-Initial
Inter-Final
Inter-No focus

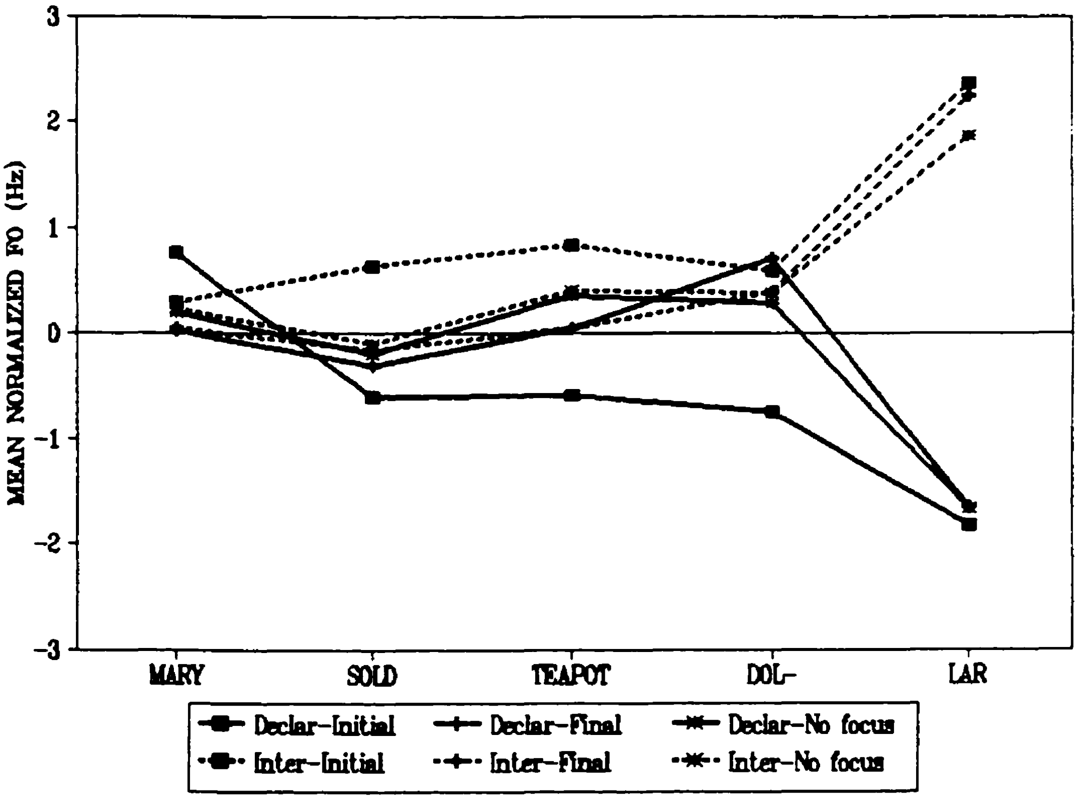


(e) SHORT HAPPY
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEA-
POT
(f) LONG HAPPY
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEAPOT
DOL-
LAR
Declar-Initial
Declar-Final
Declar-No focus
Inter-Initial
Inter-Final
Inter-No focus

(g) SHORT ANGRY



(h) LONG ANGRY



Finally, sentence terminals (K4 for short, K5 for long) exhibited a significant decrement in F_0 for declaratives and a significant increment in F_0 for interrogatives, across stimulus types. Declarative intonation with sentence-initial focus occasionally served as an exception to this pattern; for these contours, F_0 was already relatively low preceding the terminal, resulting in non-significant terminal drop in F_0 for certain stimulus types (i.e., neutral and sad stimuli, both short and long--see Figures 9a-d).

Further exploration of the MODALITY X FOCUS X KEYWORD interaction considered how differences in focus location were reflected in F_0 differences at assigned keyword positions within declarative and interrogative contours. For declarative intonation, subjects tended to set initial keywords (i.e., K1) at a comparable F_0 level irrespective of focus position. Exceptionally, a significantly higher F_0 was observed on initial focused items relative to matched unfocused items when neutral declaratives were produced (short and long) and when short, happy declaratives were elicited (cf. Eady & Cooper, 1986, for data on neutral prosody). At intervening positions in declarative contours (K2 for short, K2 and K3 for long), inter-stimulus differences reflected the low F_0 of tokens with sentence-initial focus relative to the other two focus versions; by sentence-final position (K3 for short, K4 for long), this tendency for sentences with initial focus to exhibit lower F_0 than matched sentences without initial focus was significant for all affect types and for both sentence lengths. In the case of short neutral contours (which tended to exhibit the greatest number of distinctions in general), utterances without focus were also significantly lower in F_0 in sentence-final position than utterances with focus in this position. Terminal values did not differ as a function of focus placement for the declarative utterances.

For interrogative contours, with one exception, significant differences in mean F_0 did not distinguish the three focus versions at any keyword position, for either the short or long utterances. The sole exception to this pattern was again present in the data for short, neutral utterances; for this data set, interrogatives with sentence-initial emphasis demonstrated a significant post-focus rise in F_0, rendering these utterances higher in F_0 at the second keyword position (K2) when compared to utterances with sentence-final emphasis. Otherwise, interrogative contours elicited for the three focus types were very similar in F_0 characteristics.

Finally, declaratives and interrogatives matched for location of emphasis placement were differentiated primarily at the utterance terminal, where interrogatives were always significantly higher in F_0 than declaratives. Declarative and interrogative sentences with sentence-initial focus were further differentiated at each keyword position following emphasis production (interrogatives were always higher in F_0 throughout), rendering these contours highly distinct.

Neutral (Non-affective) Stimuli

Mean F_0 values derived at each keyword position within short and long utterances spoken in a non-affective context are provided in Table 8, by group,

Table 8. Neutral Stimuli - Mean normalized F_0 values for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

SHORT		K1		K2		K3		K4		K5	
		MARY		SOLD		TEA-		POT		--	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)		
NC	Initial	0.42	-0.58	-1.71	0.96	-1.81	0.85	-2.59	3.13	--	--
	Final	-0.72	-0.44	-0.73	-0.44	1.80	0.42	-2.22	2.98	--	--
	None	-0.86	-0.63	-1.03	-0.17	-0.10	0.02	-1.84	2.97	--	--
RHD	Initial	0.58	-0.13	-1.53	0.40	-1.80	0.83	-2.46	2.77	--	--
	Final	-0.37	-0.28	-0.61	-0.12	1.47	0.82	-2.18	1.89	--	--
	None	-0.14	0.06	-0.85	0.42	-0.65	1.40	-1.86	2.23	--	--
LONG		MARY		SOLD		TEAPOT		DOL-		LAR	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
NC	Initial	1.09	-0.58	-1.80	0.60	-1.49	0.79	-1.84	0.36	-2.31	2.48
	Final	-0.54	-0.56	-0.76	-0.33	-0.63	-0.42	0.21	-0.58	-2.12	2.54
	None	-0.56	-0.74	-0.68	-0.02	-0.64	-0.17	-0.43	-0.97	-1.87	2.69
RHD	Initial	0.37	0.13	-0.80	0.25	-1.00	0.43	-1.24	0.06	-1.89	1.94
	Final	-0.16	-0.35	-0.45	-0.04	-0.08	-0.01	0.38	0.24	-2.23	1.91
	None	-0.44	-0.25	-0.49	-0.01	-0.16	0.26	-0.59	-0.02	-2.02	1.17

Note: G=Group, (.)=declarative, (?)=interrogative. Inter-subject differences in F_0 range were normalized by subtracting the mean F_0 of *all* utterances produced by the speaker from each keyword value, divided by the standard deviation.

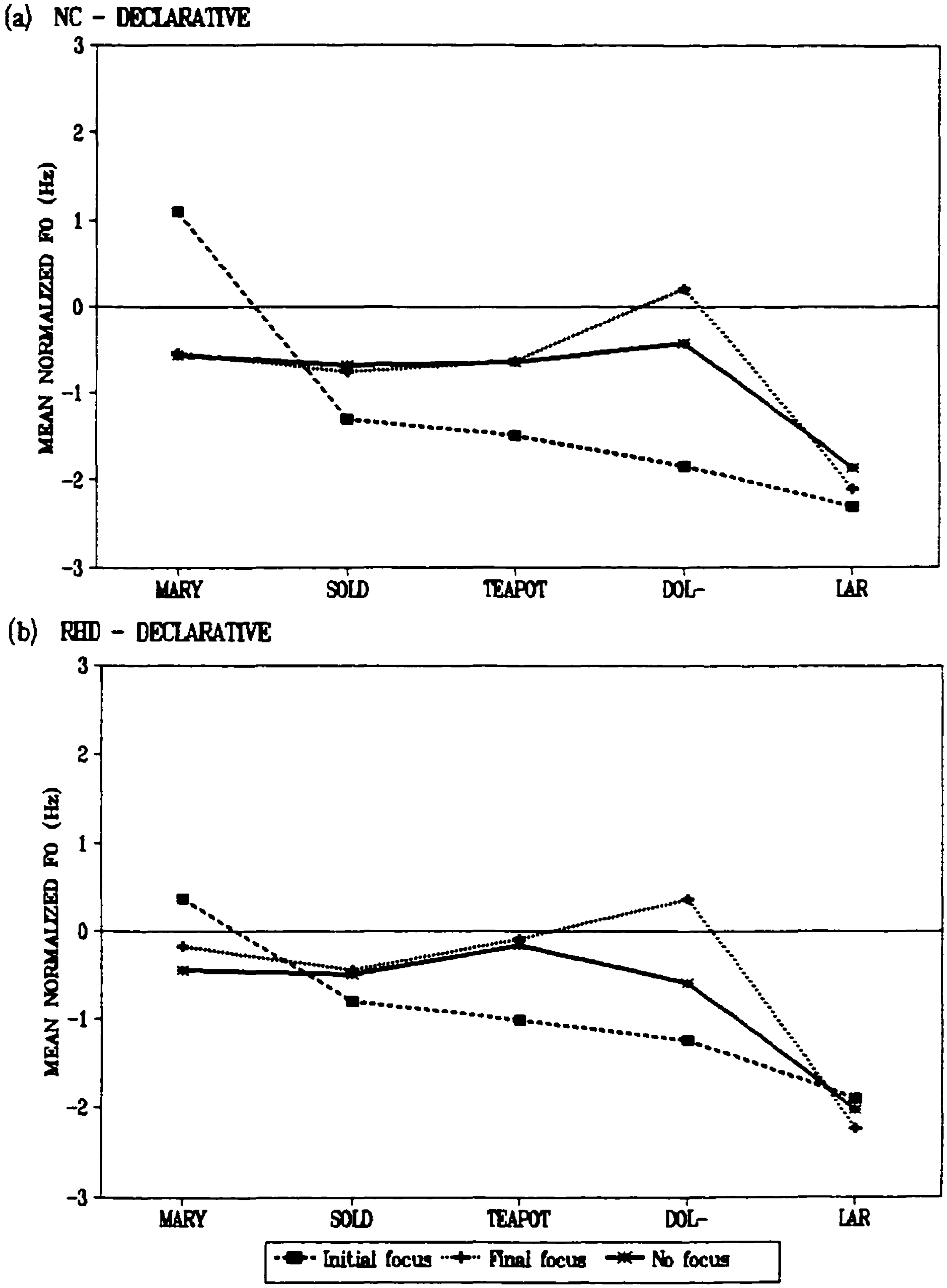
sentence modality, and focus location. Separate (GROUP X MODALITY X FOCUS X KEYWORD) ANOVAsperformed on the short and long stimuli each produced the "common" effects described for all stimuli above (and reported in

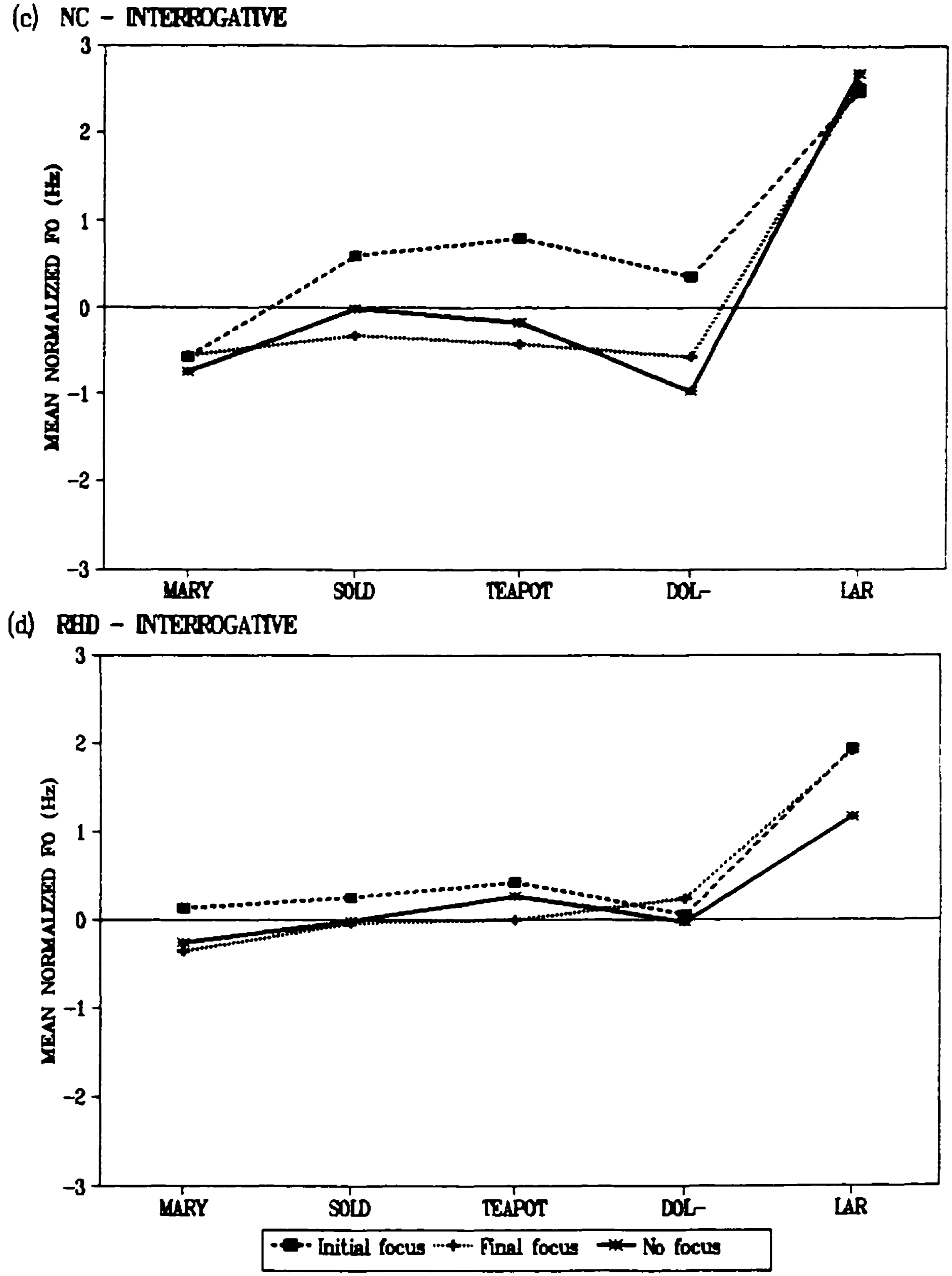
Appendix B) as well as effects unique to the neutral stimuli, but contributing to these previously-described effects (main effects: FOCUS [$F_{SHORT}(2,36) = 3.89$, $p<.05$; KEYWORD [$F_{SHORT}(3,54) = 5.70$, $p<.01$]. In addition, when long utterances were elicited in a non-affective context, significant interactions of GROUP X MODALITY X KEYWORD [$F(4,72)=3.55$, $p<.05$] and GROUP X MODALITY X FOCUS X KEYWORD [$F(8,144)=3.61$, $p=.001$] were uncovered.

Post hoc comparisons performed on the four-way interaction (displayed in Figure 10 (a-d), independently by group and modality type) looked at group differences in modulating emphasis and modality distinctions in long neutral utterances as a function of keyword position. It was revealed that, contrary to the NC subjects, the RHD patients failed to use F_0 in sentence-initial position to differentiate among declarative utterances varying in focus location: normal subjects set focused items at a significantly higher F_0 in initial position for the neutral stimuli, whereas RHD patients set all three focus versions at a uniform F_0 level. (Note, however, that both groups displayed a significant post-focus drop in F_0--compare Figures 10a and 10b).

Relatedly, in producing interrogative contours (Figures 10c-d), the RHD patients did not mark sentence-initial focus with a significant F_0 rise following the focused item, contrary to the NC subjects. Moreover, in sentence-final position, interrogative contours produced by the NC subjects demonstrated distinct patterns as a function of focus position (F_0 was significantly higher following initial focus than when focus was lacking), whereas interrogatives elicited from the RHD

Figure 10. Interaction of group, sentence modality, focus, and keyword position for long stimuli spoken by the NC and RHD subjects as declaratives (a-b) and interrogatives (c-d).




(c) NC - INTERROGATIVE
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEAPOT
DOL-
LAR
(d) RHD - INTERROGATIVE
MEAN NORMALIZED F0 (Hz)
3
2
1
0
-1
-2
-3
MARY
SOLD
TEAPOT
DOL-
LAR
Initial focus
Final focus
No focus

patients displayed comparable F_0 characteristics across focus versions in sentence-final position. These differences may have reduced the contrastiveness of the stimuli elicited from the RHD subjects as a function of mean F_0, particularly when focus was located sentence-initially. Overall, mean F_0 did not vary significantly between the two groups at a given keyword position, across stimulus types.

Sad Stimuli

The keyword F_0 data for the short and long utterances spoken in a sad context are presented in Table 9, by group, sentence modality, and focus location. Unique to the production of sad prosody, interactive effects of GROUP X FOCUS [$F(2,36) = 5.73, p<.01$] and GROUP X MODALITY X KEYWORD [$F(3,54) = 3.95, p<.05$] were found in the analysis of the short stimuli only. No significant effects unique to the production of long stimuli in a sad context emerged from this analysis.

Examination of the interaction between group and focus location indicated that when producing short utterances in a sad tone, the mean (overall) F_0 of RHD patients was significantly higher than that of NC subjects when sentence-final emphasis was present. Moreover, the RHD group produced sad sentences with final focus with a significantly higher overall F_0 than the other two focus versions, although the NC group did not differentiate the three focus types as a function of mean F_0.

Table 9. Sad Stimuli - Mean normalized F_0 values for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

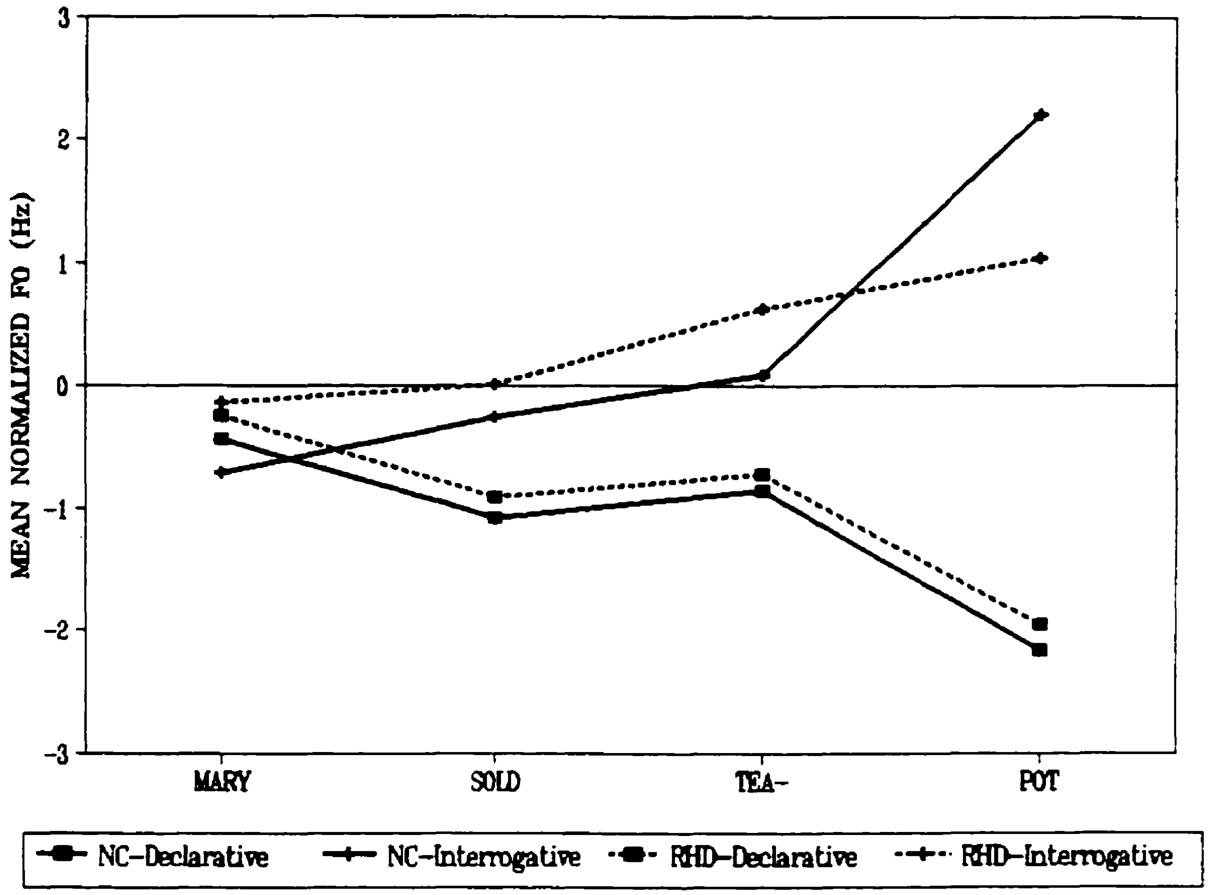
		K1		K2		K3		K4		K5	
SHORT		MARY		SOLD		TEA-		POT		--	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)		
NC	Initial	-0.47	-0.68	-1.60	0.30	-1.67	0.49	-2.07	2.68	--	--
	Final	-0.42	-0.73	-0.82	-0.87	-0.21	-0.16	-2.13	1.70	--	--
	None	-0.40	-0.70	-0.83	-0.18	-0.66	-0.08	-2.29	2.24	--	--
RHD	Initial	-0.01	-0.05	-1.50	0.20	-1.73	0.62	-2.30	0.98	--	--
	Final	-0.21	-0.21	-0.37	-0.12	0.38	0.73	-1.50	1.23	--	--
	None	-0.50	-0.14	-0.88	-0.42	-0.80	0.50	-2.07	0.90	--	--

LONG		MARY		SOLD		TEAPOT		DOL-		LAR	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
NC	Initial	-0.13	-0.57	-1.29	-0.02	-1.27	0.15	-1.57	-0.42	-2.05	1.52
	Final	-0.42	-0.45	-1.05	-0.66	-0.89	-0.51	-0.50	-0.69	-1.92	1.39
	None	-0.56	-0.39	-1.00	-0.46	-0.76	-0.56	-0.79	-0.94	-1.55	1.70
RHD	Initial	0.15	-0.12	-1.16	-0.23	-1.14	0.19	-1.14	0.16	-2.05	1.54
	Final	-0.35	-0.38	-0.55	-0.42	-0.27	-0.24	-0.35	-0.27	-2.09	1.14
	None	-0.32	-0.29	-0.65	-0.09	-0.16	0.34	-0.50	-0.12	-1.26	1.83

Note: G=Group, (.)=declarative, (?)=interrogative. Inter-subject differences in F_0 range were normalized by subtracting the mean F_0 of *all* utterances produced by the speaker from each keyword value, divided by the standard deviation.

Exploration of the interaction among group, modality, and keyword F_0, displayed in Figure 11, indicated that the RHD group marked interrogation with a significantly lower F_0 on the utterance terminal (K4) than the NC subjects when

Figure 11. Interaction of group, sentence modality, and keyword position for short stimuli spoken in a sad context.



speaking in a sad tone. In addition, the F_0 of the RHD patients was significantly higher in sentence-final position (K3) for interrogatives than declaratives, contrary to the normal pattern, for which declarative/ interrogative distinctions were produced only at the utterance terminal.

Happy Stimuli

Table 10 summarizes the keyword F_0 data collected for short and long utterances spoken in a happy context for each subject group, as a function of sentence modality and focus distribution. Significant effects specific to the production of happy prosody did not emerge for the long utterances. However, for short utterances, significant main effects were found for GROUP [$F(1,18) = 7.96, p=.01$] and KEYWORD [$F(3,54) = 6.67, p=.001$], and significant interactions were found for GROUP X MODALITY X KEYWORD [$F(3,54) = 5.79, p<.01$] and GROUP X MODALITY X FOCUS X KEYWORD [$F(6,108) = 2.45, p<.05$]. Figure 12 (a-b) illustrates the four-way interaction graphically, by sentence modality.

Inspection of these data indicated that when producing declarative utterances in a happy tone, RHD patients again did not differentiate among the three focus versions in sentence-initial position (NC subjects produced significantly higher values on focused items). Moreover, the RHD patients did not raise F_0 significantly to signal focus on sentence-final items, contrary to the NC subjects (see Figure 12a). When producing interrogative intonation in a happy tone, RHD

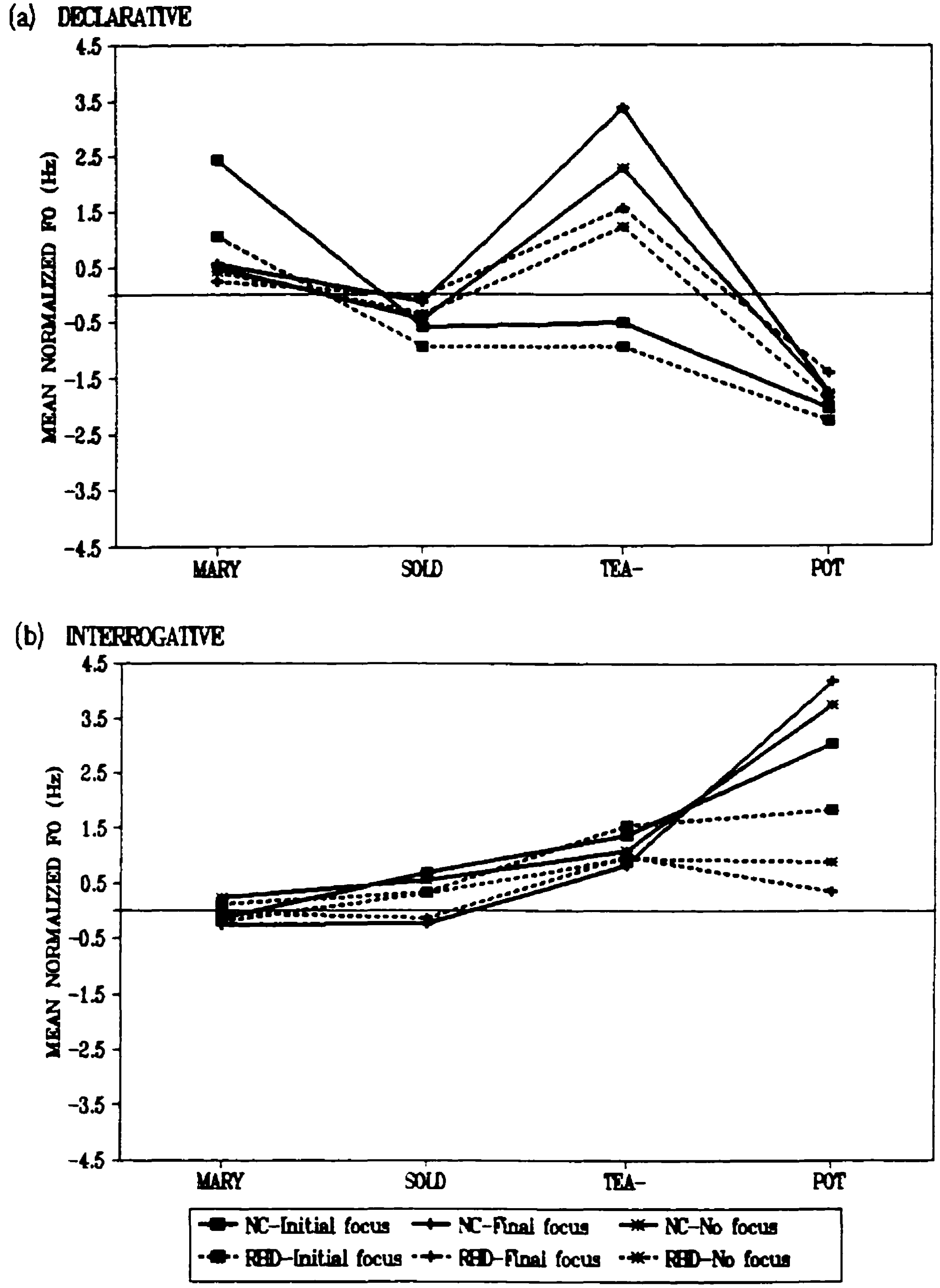
Table 10. Happy Stimuli - Mean normalized F_0 values for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

		K1		K2		K3		K4		K5	
SHORT		MARY		SOLD		TEA-		POT		--	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)		
NC	Initial	2.41	-0.14	-0.58	0.70	-0.52	1.37	-2.06	3.05	--	--
	Final	0.54	-0.25	-0.13	-0.24	3.38	0.83	-1.78	4.21	--	--
	None	0.51	0.23	-0.46	0.56	2.29	1.10	-1.79	3.75	--	--
RHD	Initial	1.05	0.12	-0.93	0.34	-0.94	1.55	-2.27	1.84	--	--
	Final	0.23	-0.03	-0.03	-0.14	1.56	0.98	-1.42	0.35	--	--
	None	0.40	-0.19	-0.35	0.32	1.22	0.94	-1.94	0.89	--	--
LONG		MARY		SOLD		TEAPOT		DOL-		LAR	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
NC	Initial	2.26	0.58	-0.68	0.57	-0.48	0.93	-0.15	0.59	-2.27	2.80
	Final	0.19	0	-0.39	-0.24	-0.01	-0.31	2.41	0.03	-1.70	3.24
	None	1.21	0.13	-0.24	0.07	0.64	0.08	1.66	0	-1.83	2.64
RHD	Initial	0.84	0.59	-0.29	0.51	-0.29	0.67	-0.23	0.51	-2.16	1.64
	Final	0.50	-0.29	0.17	-0.19	0.77	0.18	1.60	0.35	-1.98	2.18
	None	0.25	-0.44	0.16	0.05	0.81	0.12	0.85	0.12	-2.05	1.34

Note: G=Group, (.)=declarative, (?)=interrogative. Inter-subject differences in F_0 range were normalized by subtracting the mean F_0 of *all* utterances produced by the speaker from each keyword value, divided by the standard deviation.

patients exhibited a significantly lower terminal F_0 than NC subjects for two of the three focus types (no focus and sentence-final focus). Group differences in terminal measures appeared to emerge as a result of the RHD patients' failure to

Figure 12. Interaction of group, sentence modality, focus, and keyword position for short (a) declaratives and (b) interrogatives spoken in a happy context.



produce a significant rise in F_0 in this position for any of the three focus versions, as measured between K3 and K4 (see Figure 12b).

Angry Stimuli

Keyword F_0 values computed for the short and long sentences elicited in an angry tone are supplied in Table 11, by group, sentence modality, and focus type. Unique to the production of angry prosody, significant main effects of KEYWORD [$F(3,54) = 4.15, p=.01$] and FOCUS [$F(2,36) = 12.05, p<.001$], and a significant interaction of GROUP X FOCUS [$F(2,36) = 4.08, p<.05$] emerged for the short utterances. Exploration of the interaction revealed that the RHD group produced angry utterances with sentence-initial focus with a significantly lower F_0, overall, than utterances with sentence-final or no focus. The NC group did not differentiate the three focus types as a function of mean F_0, when considered overall. No differences in mean F_0 were present between-groups for utterances matched in focus location.

For only the long stimuli, a significant GROUP X MODALITY interaction emerged [$F(1,18)=8.82, p<.01$]. Comparison of the marginal means indicated that both groups produced declaratives with a significantly lower mean F_0 than interrogatives, overall, but that declaratives produced by the RHD patients exhibited a significantly higher F_0 than those produced by the NC subjects (interrogatives did not differ significantly for the two groups).

Table 11. Angry Stimuli - Mean normalized F_0 values for the NC and RHD groups as a function of sentence modality and focus location (by sentence length).

		K1		K2		K3		K4		K5	
SHORT		MARY		SOLD		TEA-		POT		--	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)		
NC	Initial	0.32	0.08	-1.39	0.55	-1.06	0.94	-2.12	2.85	--	--
	Final	0.19	-0.50	-0.45	-0.64	1.46	0.90	-2.29	2.93	--	--
	None	0.03	0.30	-0.29	0.02	0.36	1.14	-2.16	2.46	--	--
RHD	Initial	0.72	0.26	-1.21	-0.09	-1.43	0.30	-2.29	1.03	--	--
	Final	0.37	-0.03	-0.36	0.05	1.50	0.99	-2.10	1.56	--	--
	None	0.87	0.26	0.32	0.16	0.89	1.32	-1.56	1.66	--	--
LONG		MARY		SOLD		TEAPOT		DOL-		LAR	
G	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
NC	Initial	0.46	0.32	-0.79	0.66	-0.90	1.03	-0.98	0.42	-1.96	2.58
	Final	-0.07	0.16	-0.45	-0.32	-0.09	0	0.69	0.41	-1.90	2.58
	None	-0.08	0.39	-0.44	-0.38	0.10	0.41	0.23	0.40	-2.40	2.16
RHD	Initial	1.06	0.27	-0.41	0.61	-0.26	0.66	-0.53	0.75	-1.70	2.17
	Final	0.13	-0.07	-0.18	0.02	0.18	0.09	0.73	0.39	-1.40	1.91
	None	0.47	0.07	0.04	0.18	0.60	0.41	0.34	0.35	-0.93	1.58

Note: G=Group, (.)=declarative, (?)=interrogative. Inter-subject differences in F_0 range were normalized by subtracting the mean F_0 of *all* utterances produced by the speaker from each keyword value, divided by the standard deviation.

Summary of the Keyword F_0 Data

In brief, analysis of each group's ability to modulate mean F_0 on designated keywords within short and long sentences uncovered prosodic irregularities in the utterances produced by the RHD patients. Generally, both RHD and NC subjects produced patterns that resembled each other in overall shape; however, the RHD patients tended to mark focus and (at times) modality distinctions less consistently and with smaller F_0 excursions than the NC subjects. Moreover, the RHD group produced fewer F_0 distinctions among contours differing in focus location than the normal group when viewed across stimulus types. Generally, the modulation of mean F_0 throughout an utterance was shown to depend highly on the interaction between sentence position, focus location, and sentence modality in speech production, irrespective of affective tone or sentence length.

(iii) F_0 Accent

To characterize the relative *magnitude* of F_0 change associated with content words with and without focus (i.e., to identify the presence of a local "F_0 accent"), F_0 range values within sentence-initial and sentence-final keywords (i.e., the range in F_0 from first to second syllable within the keyword) were computed in both short and long sentences. Table 12 presents the normalized data for the two diagnostic groups as a function of emotional prosody, sentence modality, and emphasis location within the utterance.

Table 12. Extent of F_0 change ("F_0 Accent") associated with emphasis on sentence-initial (K1) and sentence-final (K3/K4) keywords produced by the NC and RHD subjects, as a function of emotional tone, sentence modality, and focus location (by sentence length).

		NC						RHD					
SHORT		Initial		Final		No focus		Initial		Final		No focus	
Emo	*Mod*	K1	K3	K1	K3	K1	K3	K1	K3	K1	K3	K1	K3
N	(.)	0.43	0.19	0.14	0.84	0.10	0.46	0.40	0.15	0.11	0.64	0.18	0.31
	(?)	0.30	0.37	0.08	0.59	0.13	0.65	0.24	0.35	0.10	0.53	0.12	0.41
S	(.)	0.26	0.16	0.17	0.42	0.17	0.37	0.31	0.14	0.13	0.39	0.11	0.24
	(?)	0.25	0.38	0.16	0.49	0.12	0.46	0.14	0.31	0.12	0.34	0.12	0.33
H	(.)	0.53	0.27	0.20	0.89	0.22	0.74	0.39	0.23	0.21	0.58	0.20	0.50
	(?)	0.33	0.53	0.12	0.73	0.21	0.65	0.26	0.50	0.12	0.43	0.12	0.44
A	(.)	0.48	0.27	0.17	0.74	0.21	0.52	0.39	0.18	0.16	0.58	0.24	0.35
	(?)	0.30	0.35	0.09	0.58	0.21	0.59	0.22	0.26	0.10	0.52	0.17	0.42
LONG		K1	K4	K1	K4	K1	K4	K1	K4	K1	K4	K1	K4
N	(.)	051	0.12	0.14	0.51	0.12	0.34	0.22	0.17	0.18	0.46	0.12	0.34
	(?)	0.34	0.42	0.15	0.73	0.18	0.70	0.19	0.39	0.17	0.55	0.10	0.50
S	(.)	0.24	0.15	0.16	0.41	0.15	0.27	0.24	0.16	0.17	0.34	0.12	0.26
	(?)	0.29	0.42	0.14	0.51	0.15	0.60	0.21	0.30	0.07	0.34	0.10	0.44
H	(.)	0.60	0.41	0.20	0.76	0.27	0.64	0.30	0.34	0.18	0.54	0.16	0.50
	(?)	0.41	0.56	0.21	0.83	0.19	0.75	0.21	0.46	0.08	0.55	0.11	0.59
A	(.)	0.46	0.31	0.15	0.62	0.23	0.56	0.30	0.31	0.12	0.50	0.18	0.46
	(?)	0.31	0.46	0.19	0.66	0.30	0.69	0.19	0.36	0.16	0.57	0.14	0.54

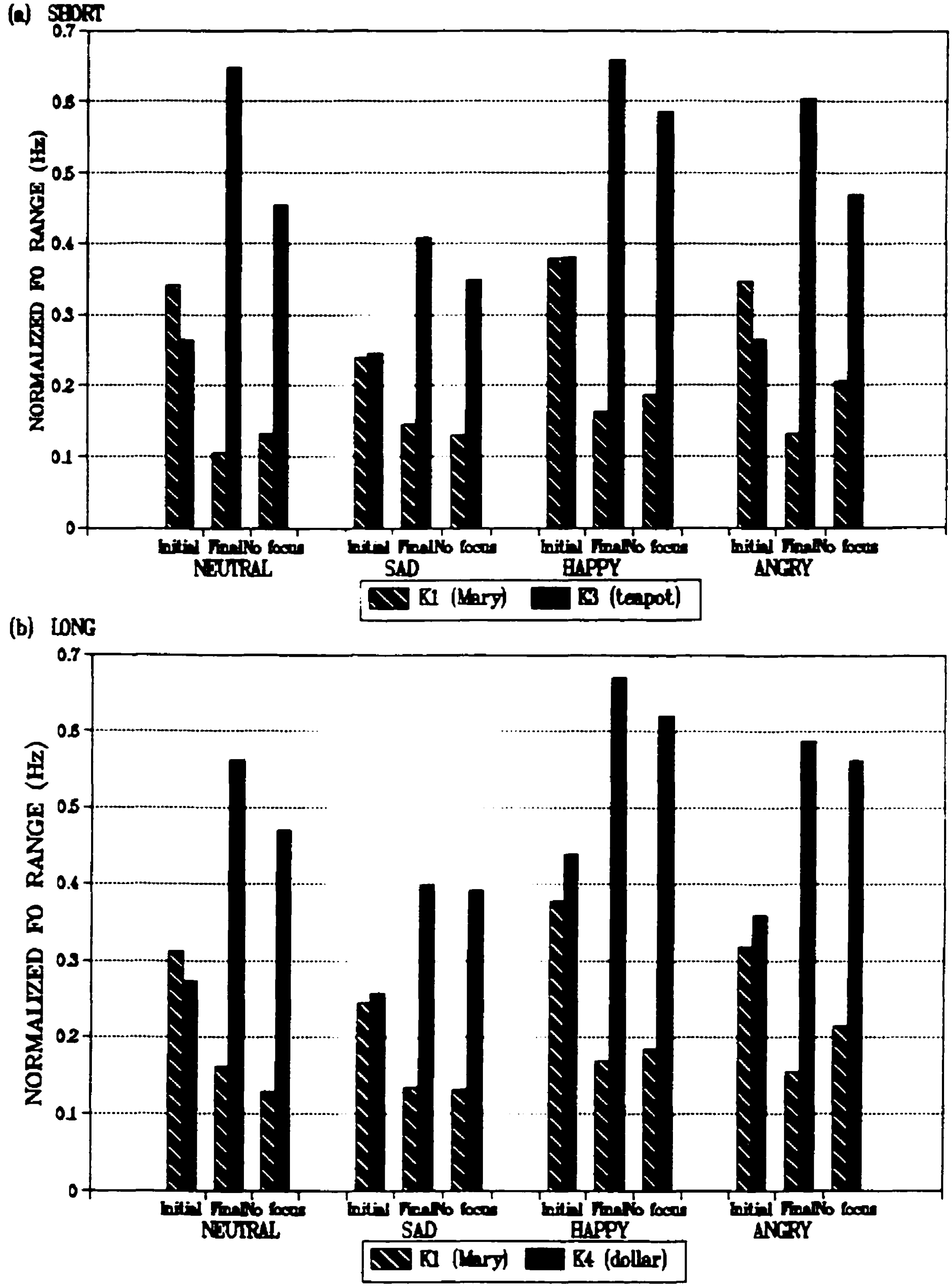
Note: N=Neutral, S=Sad, H=Happy, A=Angry, (.)=declarative, (?)=interrogative. Range values were normalized for inter-subject differences in mean F_0 by dividing the F_0 range of each keyword by the mean utterance F_0 for each speaker.

Separate five-way (GROUP X EMOTION X MODALITY X FOCUS X KEYWORD) ANOVAs performed on the short and long data each yielded a main effect for GROUP [$F_{SHORT}(1,18) = 5.97, p<.05$; $F_{LONG}(1,18) = 5.44, p<.05$]. Closer inspection of these data indicated that the RHD patients produced significantly less F_0 variation than the NC subjects on content words associated with emphatic stress in the present experiment; this finding suggests that the RHD patients were restricted in the capacity to modulate *short-term* F_0 parameters for prosodic signaling. Interactions with the group factor were not observed for the short stimuli, but did emerge for the long stimuli (see below).

For both the short and long stimuli, significant three-way interactions of EMOTION X FOCUS X KEYWORD [$F_{SHORT}(6,108) = 5.35, p<.001$; $F_{LONG}(6,108) = 3.33, p<.01$], MODALITY X FOCUS X KEYWORD [$F_{SHORT}(2,36)=25.17, p<.001$; $F_{LONG}(2,36)=5.72, p<.01$], and EMOTION X MODALITY X FOCUS [$F_{SHORT}(6,108)=2.49, p<.05$; $F_{LONG}(6,108) = 2.21, p<.05$] were witnessed. These effects are represented graphically in Figures 13, 14, and 15, respectively. Other statistically significant effects observed for the short and/or long stimuli which contributed to the three-way interactions are reported in Appendix C.

Follow-up tests performed on the EMOTION X FOCUS X KEYWORD interaction (presented in Figure 13 (a-b) by sentence length) first compared the magnitude of F_0 excursions produced on sentence-initial (K1) versus sentence-final (K3 for short, K4 for long) keywords as a function of emotion and emphasis

Figure 13. Interaction in F_0 range of emotion, focus, and keyword position for (a) short and (b) long stimuli.



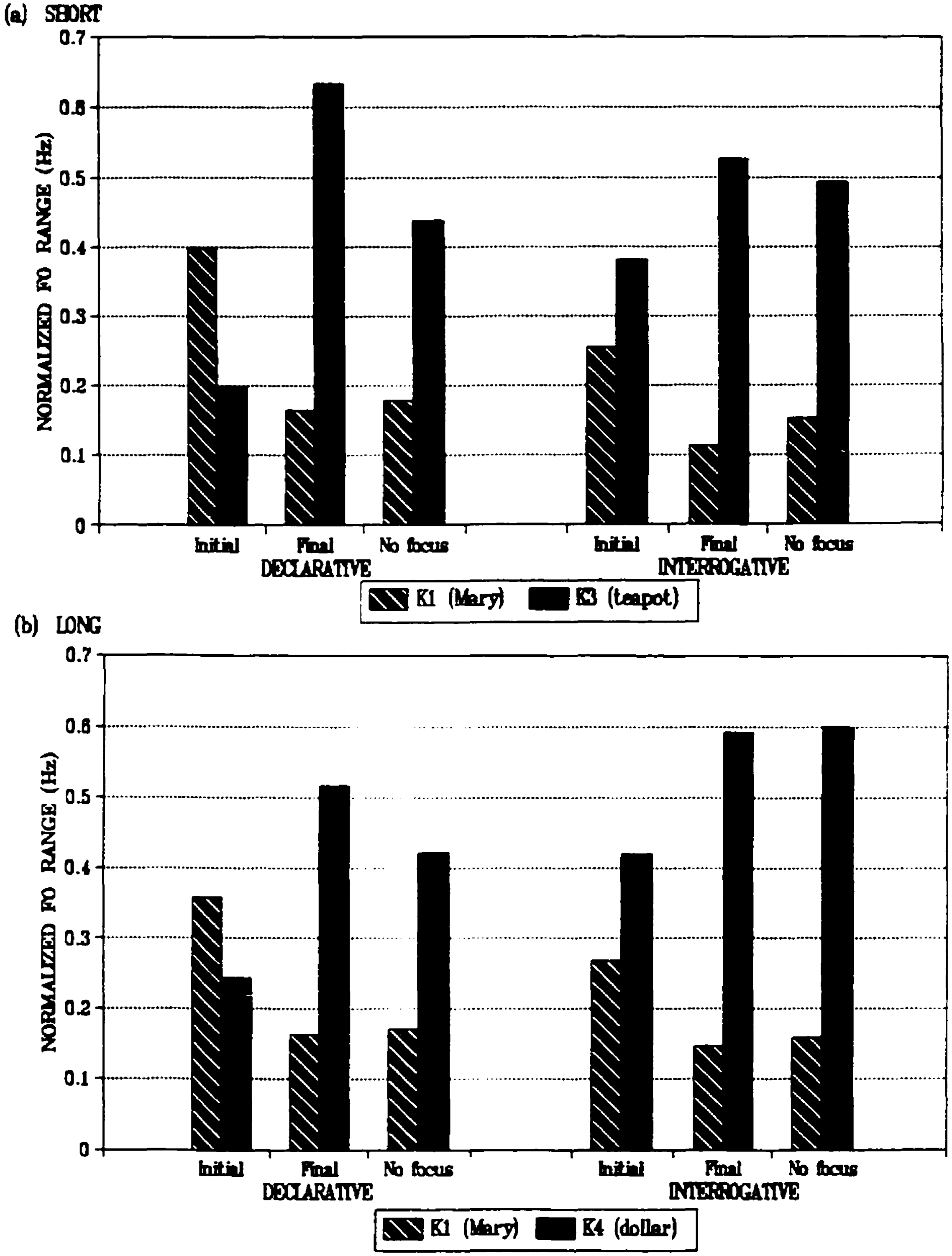
placement. For utterances without emphasis or with sentence-terminal emphasis, sentence-final items (K3/K4) displayed significantly larger F_0 excursions than sentence-initial items (K1) for all emotion types, in both short and long sentences. In contrast, utterances characterized by sentence-initial emphasis displayed comparable F_0 change at both K1 and K3/K4 for all emotion types. (This pattern contrasts with that reported in the keyword duration data, where sentence-*final* emphasis appeared to be marked by an *absence* of acoustic differences between sentence-initial and sentence-final items.)

Examination of differences in F_0 range at designated keyword positions indicated that sentence-initial items (K1) displayed greater variation when focused than when unfocused (i.e., than tokens with terminal or no focus) irrespective of the emotional tone of the speaker, as expected. Exceptionally, short sentences spoken in a sad tone were undifferentiated at K1 by F_0 range for utterances varying in emphasis placement. The magnitude of F_0 accents produced at K1 did not significantly differ as a function of the emotional tone of the utterance for any of the three focus versions. On sentence-final keywords (K3, K4), F_0 range was significantly *reduced* when focus was produced sentence-initially than when focus was absent or produced sentence-terminally, for all four emotion types. For short sentences elicited in a neutral and happy context, K3 range values were also significantly larger when focus was present in this position than when focus was absent in the utterance altogether. In contrast to sentence-initial items, the magnitude of F_0 excursions associated with sentence-final items was shown to be influenced by the emotional mode of the utterance, happy sentences displaying consistently larger range values sentence-finally than sad sentences (short and long) and neutral sentences (long only) irrespective of focus location. Angry stimuli also exhibited significantly larger F_0 excursions sentence-finally than sad stimuli when sentences were long, and when short sentences were produced with terminal emphasis. Finally, neutral stimuli displayed significantly larger F_0 ranges in sentence-final position than sad stimuli only when short and long utterances were produced with sentence-final emphasis (review Figure 13a-b).

The MODALITY X FOCUS X KEYWORD interaction is presented visually in Figure 14 (a-b), independently for the short and long stimuli. Post hoc exploration of this interaction indicated that the magnitude of F_0 variation associated with sentence-initial and sentence-final keywords was strongly influenced by sentence modality only when emphasis was produced sentence-initially; for these tokens, K1 displayed significantly larger F_0 excursions than K3 or K4 when declaratives were elicited, but the *opposite* pattern was evident (K3/K4>K1) when interrogatives were elicited. Similarly, interrogatives with initial emphasis and sentences with final or no emphasis demonstrated qualitatively similar patterns when the relative magnitude of F_0 change on select keywords was considered (see Figure 14).

For both declarative and interrogative contours, focus at K1 was again signaled by means of larger F_0 excursions than those produced in utterances without focus in that position (which did not differ). At K3 or K4, however,

Figure 14. Interaction in F_0 range of sentence modality, focus, and keyword position for (a) short and (b) long stimuli.

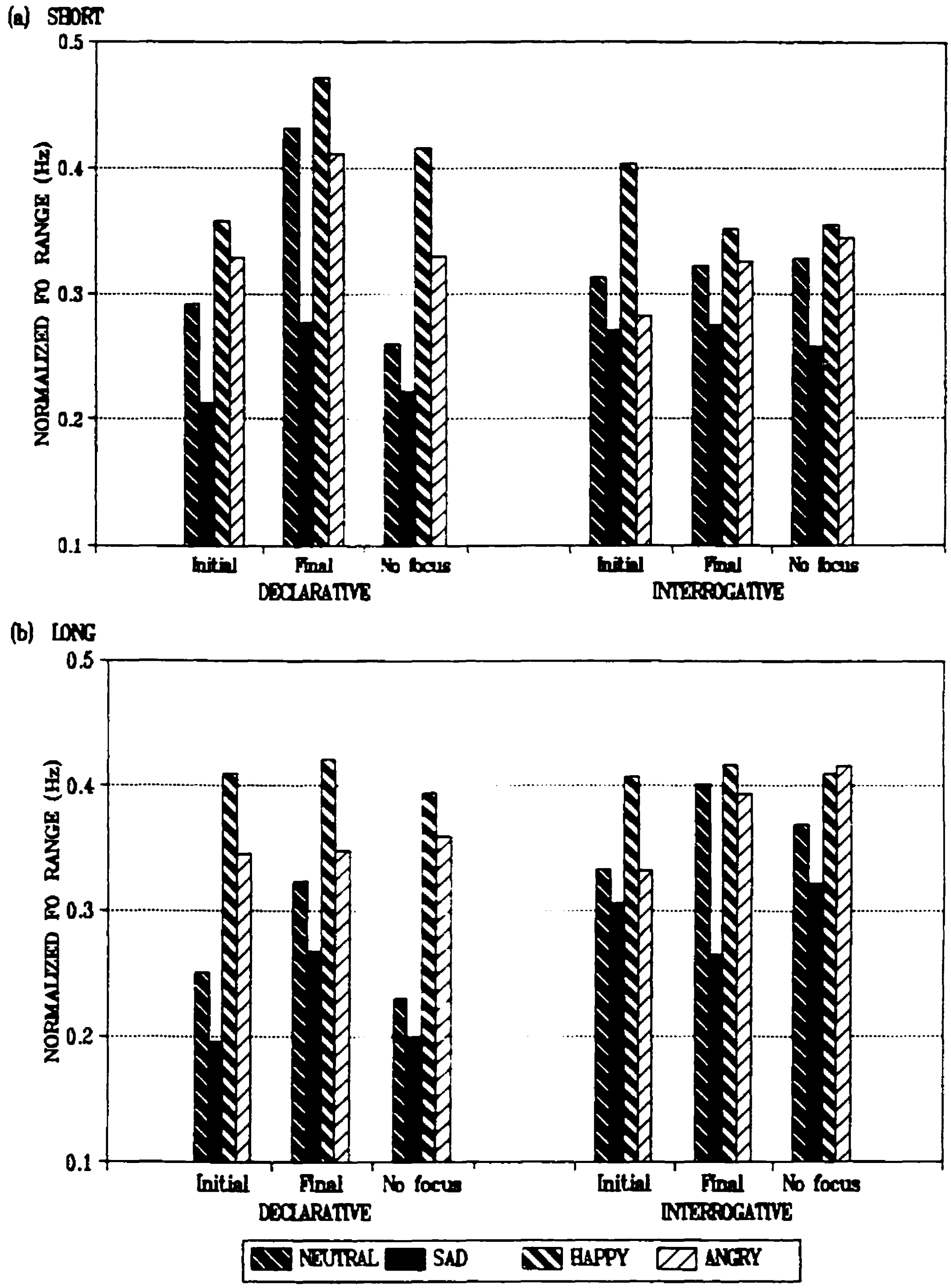


focused items only demonstrated consistently larger F_0 excursions than matched items contained in sentences with sentence-initial emphasis (which demonstrated severely restricted F_0 ranges at K3/K4). The amount of F_0 change associated with focus on sentence-final items could not be differentiated from that contained in sentences without focus when interrogatives were produced (review Figure 14a-b).

The interaction of EMOTION, MODALITY, and FOCUS is conveyed in Figure 15 (a-b), by sentence length. Closer inspection of these data suggested that local F_0 accents were of significantly greater magnitude in happy utterances than in sad utterances for the majority of stimulus types (4/6 versions for short stimuli, 5/6 versions for long stimuli). Utterances elicited in an angry tone also tended to be associated with greater F_0 change on content words than those elicited in a sad tone, a pattern true for half of the short and long tokens. Further differences among the stimuli did not point to systematic tendencies, although it is noteworthy that differences (where noted) always reflected the general relationship among the four affect types inasmuch as *long-term* measures of F_0 variation are concerned (reported in detail below). These findings suggest that the magnitude of F_0 accents underlying linguistic focus are often influenced by the emotional mode of the speaker, larger F_0 accents being associated with emotions characterized by greater long-term variation in F_0 (e.g., happy, angry).

The emergence of a significant four-way interaction (EMOTION X MODALITY X FOCUS X KEYWORD [$F(6,108)=2.38$, $p<.05$]) for only the short stimuli reinforces earlier observations (in the keyword duration and F_0 data)

Figure 15. Interaction in F_0 range of emotion, sentence modality, and focus for (a) short and (b) long stimuli.



that short intonation contours tend to display greater acoustic interaction among the independent prosodic forms when encoded concurrently. Investigation of the relationship among these four variables suggested qualitatively similar patterns to those described in detail in the discussion of three-way interactions and do not merit further comment (review Figures 13, 14, 15).

For only the long stimuli, interactive effects of GROUP X FOCUS X KEYWORD [$F(2,36)=4.26$, $p<.05$] and GROUP X EMOTION X FOCUS X KEYWORD [$F(6,108)=2.34$, $p<.05$] reached statistical significance. Inspection of the four-way interaction revealed that the proportion of F_0 change associated with K1 relative to K4 did not differ in the productions elicited from the RHD and NC subjects across stimulus types. However, for all three focus versions, the RHD patients produced smaller F_0 excursions than the normal subjects on content words contained in happy sentences (as well as one version of neutral sentences). This finding suggests that happy prosody may have posed a particular challenge to the RHD speakers, affecting the normal extent to which F_0 variation could be employed to signal emphasis at the word level for these stimuli. Finally, at designated keyword positions, the RHD patients again demonstrated a tendency to produce fewer distinctions among stimuli differing in focus location than the normal subjects. This trend was especially pronounced at K1, where the RHD patients failed to produce the normal pattern of increased F_0 variation when focus occurred in this position as compared to when focus was absent or placed sentence-terminally (review Figure 14 for the overall pattern).

2) Global Measures

In Experiment 1, three measures were derived to characterize acoustic change occurring throughout short and long utterances as a whole: speech rate, mean F_0, and F_0 range. These features are believed to be especially pertinent to the expression of emotional prosody (e.g., Bolinger, 1978; Scherer, 1986). Accordingly, emotion type was considered directly within the analyses performed on the three "global" acoustic measures; for each (speech rate, mean F_0, F_0 range), separate 2 X 4 X 2 X 3 mixed ANOVAs in a full factorial design were performed on the data derived for short and long utterances. Group membership (NC, RHD) constituted the between-groups factor, and EMOTION (neutral, sad, happy, angry), MODALITY (declarative, interrogative), and FOCUS (none, initial, final) served as the within-subjects variables for each analysis. As before, effects common to both the short and long stimuli are reported first for each of the three measures, followed by a description of effects attributable to only the short or long stimuli in each case.

(i) Speech Rate

To establish whether differences in rate of articulation were consistent with distinctions in prosodic content, the number of syllables contained in each short and long utterance was divided by the total utterance duration to arrive at an indication of syllables spoken per second. These data are illustrated in Table 13 for each sentence length, as a function of emotion type, sentence modality, and focus location (by group).

Table 13. Mean speech rate (syllables/second) for the NC and RHD groups as a function of emotion type, sentence modality, and focus location (by sentence length).

			NC				RHD			
Length	*Mod*	*Focus*	N	S	H	A	N	S	H	A
Short	Declar	Initial	4.80	3.89	4.54	4.33	4.64	4.06	4.33	4.27
		Final	4.85	3.57	4.40	4.06	4.68	3.99	4.28	4.15
		None	4.97	3.93	4.68	4.01	4.45	4.07	4.38	4.24
	Inter	Initial	4.50	3.93	4.27	4.36	4.10	4.13	4.23	4.10
		Final	4.43	3.62	4.35	4.19	4.34	4.09	4.41	4.25
		None	4.43	3.94	4.55	4.20	4.36	4.16	4.46	4.24
Long	Declar	Initial	5.55	4.44	5.02	4.79	5.36	4.82	5.29	5.05
		Final	5.17	4.25	5.03	4.71	5.19	4.85	5.11	5.04
		None	5.46	4.57	5.02	4.73	5.09	4.87	5.13	4.92
	Inter	Initial	5.16	4.54	5.08	5.00	4.92	4.57	4.96	4.78
		Final	5.15	4.46	4.95	4.69	5.06	4.86	5.17	4.99
		None	5.30	4.67	5.22	4.52	5.32	4.93	5.04	4.98

Note: Declar= Declarative, Inter=Interrogative, N=Neutral, S=Sad, H=Happy, A=Angry.

The ANOVAs performed on the speech rate data for short and long sentences yielded several common effects. Namely, a significant main effect of EMOTION [$F_{SHORT}(3,54)=44.92, p<.001$: $F_{LONG}(3,54)=41.12, p<.001$] and a significant interactive effect of EMOTION X FOCUS [$F_{SHORT}(6,108)=3.05$, $p<.01$; $F_{LONG}(6,108)=2.66$, $p<.05$] emerged in both analyses. Post hoc inspection of the interaction (presented in Figure 16) indicated that for both the short and long stimuli, neutral and happy prosody (which did not significantly differ) were produced at a significantly faster rate than sad prosody, and neutral prosody was spoken significantly faster than angry prosody, irrespective of focus location within the utterance. Happy sentences were also significantly faster than angry sentences when utterances lacked focus, and angry sentences were faster than sad sentences for two of the three focus conditions (sentence-initial and sentence-final focus). Speaking rate of utterances spoken in a given affective mode did not differ significantly as a function of emphasis location in any case.

A further interaction between GROUP and EMOTION [$F_{SHORT}(3,54)=7.80$, $p<.001$; $F_{LONG}(3,54)=6.64$, $p=.001$] is illustrated in Figure 17. Follow-up tests established that the RHD patients produced sad prosody at a significantly higher rate than the NC subjects, for both short and long sentences. No other group differences emerged as a function of emotion type. Examination of each group's ability to use speech rate to differentiate the four emotions suggested that for the NC group, neutral and happy prosody were both produced significantly faster than angry prosody, which in turn was produced significantly faster than sad prosody. For the RHD group, however, fewer distinctions were evident; neutral and happy prosody were produced significantly faster than sad

Figure 16. Interaction of emotion and focus for the short and long stimuli.

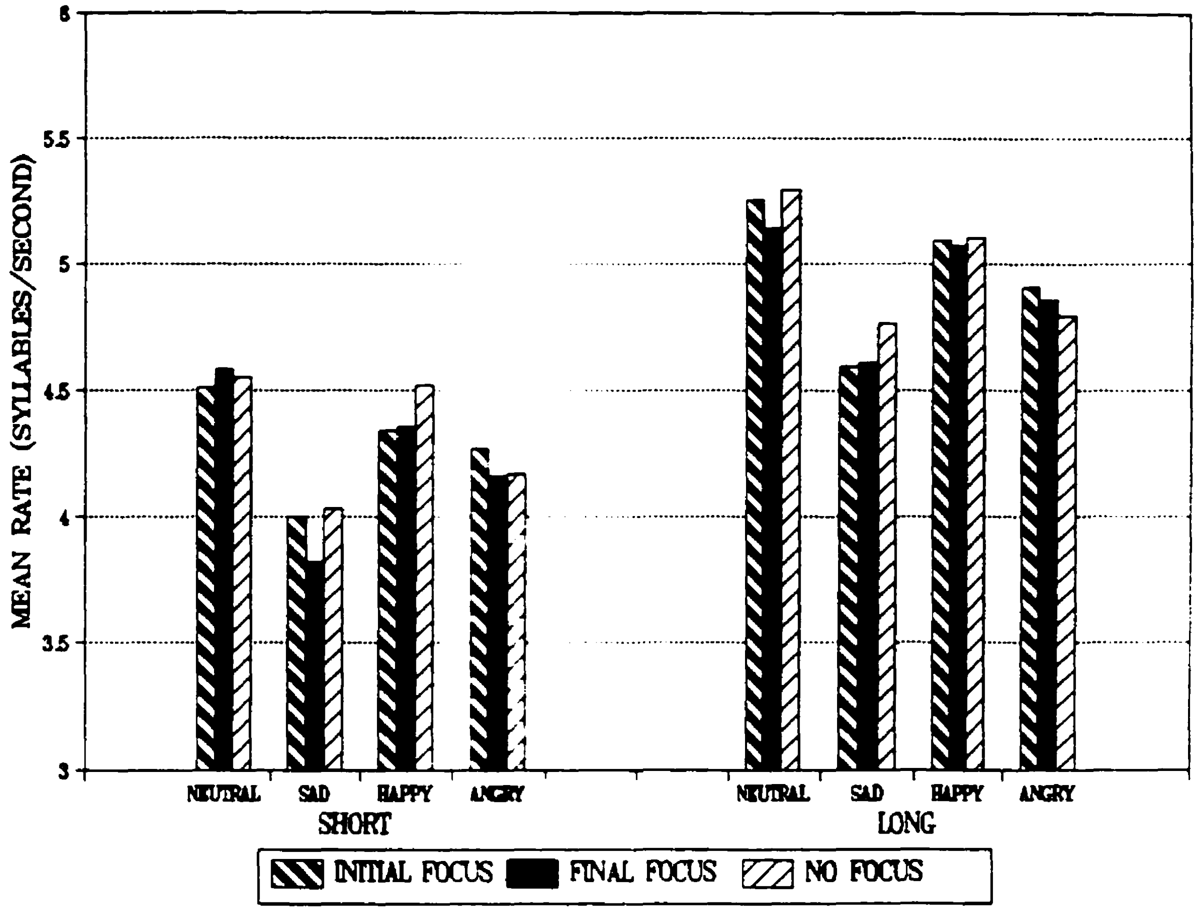
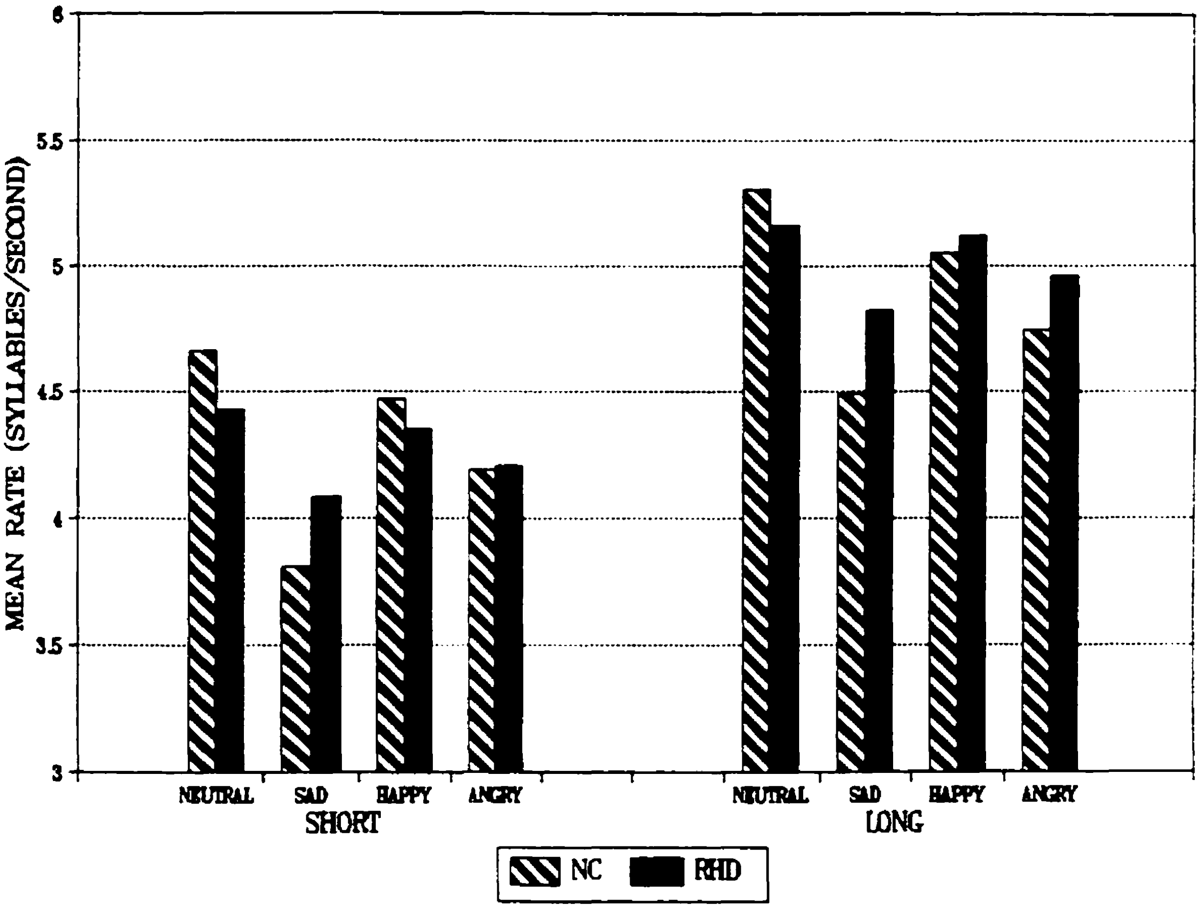


Figure 17. Interaction of group and emotion for the short and long stimuli.



prosody, but angry prosody did not significantly differ in rate from either neutral, happy, or sad prosody.

For only the short stimuli, an EMOTION X MODALITY interaction was further observed [$F(3,54)=12.42$, $p<.001$]. Post hoc comparisons indicated that for declaratives collapsed across groups, all four emotions could be distinguished in terms of speech rate; neutral was produced significantly faster than happy, happy exceeded angry in rate, and angry exceeded sad in rate. For interrogatives, significant rate differences were evident only between sad prosody (slowest) and neutral, happy, and angry prosody (which did not significantly differ). Only neutral prosody differed in rate as a function of linguistic modality, neutral declaratives being produced significantly faster than neutral interrogatives.

Finally, for only the long stimuli, significant interactions of MODALITY X FOCUS [$F(2,36)=4.54$, $p<.05$] and GROUP X MODALITY X FOCUS [$F(2,36)=4.57$, $p<.05$] were noted. Inspection of the three-way interaction yielded few significant comparisons of note; RHD patients produced declarative utterances with sentence-final focus significantly faster than NC subjects, and for the RHD patients only, sentences with initial focus were spoken significantly faster as declaratives than as interrogatives.

(ii) Mean F_0

Mean (normalized) F_0 values computed for the short and long utterances elicited from each subject group are presented in Table 14, as a function of

Table 14. Mean F_0 of short and long utterances produced by the NC and RHD groups as a function of emotion type, sentence modality, and focus location.

			NC				RHD			
Length	*Mod*	*Focus*	N	S	H	A	N	S	H	A
Short	Declar	Initial	-1.42	-1.45	-0.18	-1.06	-1.30	-1.39	-0.77	-1.05
		Final	-0.46	-0.89	0.52	-0.23	-0.42	-0.42	0.09	-0.15
		None	-0.95	-1.04	0.14	-0.51	-0.87	-1.06	-0.16	0.13
	Inter	Initial	1.09	0.70	1.25	1.11	0.97	0.44	1.00	0.38
		Final	0.63	-0.02	1.14	0.68	0.58	0.41	0.29	0.65
		None	0.55	0.32	1.41	0.98	1.04	0.31	0.49	0.86
Long	Declar	Initial	-1.17	-1.26	-0.26	-0.83	-0.91	-1.11	-0.42	-0.36
		Final	-0.76	-0.96	0.11	-0.34	-0.51	-0.73	0.21	-0.11
		None	-0.84	-0.93	0.29	-0.52	-0.74	-0.58	0.01	0.11
	Inter	Initial	0.73	0.09	1.09	1.00	0.56	0.31	0.79	0.90
		Final	0.13	-0.18	0.54	0.57	0.35	-0.04	0.45	0.47
		None	0.16	-0.13	0.57	0.59	0.23	0.33	0.24	0.51

Note: Declar= Declarative, Inter=Interrogative, N=Neutral, S=Sad, H=Happy, A=Angry. Inter-subject differences in F_0 range were normalized by subtracting the mean F_0 of *all* utterances produced by the speaker from each value, divided by the standard deviation.

emotion type, sentence modality, and emphatic stress location. Four-way (GROUP X EMOTION X MODALITY X FOCUS) ANOVAs, performed separately on the short and long data, each yielded significant main effects for EMOTION [$F_{SHORT}(3,54)=13.82$, $p<.001$; $F_{LONG}(3,54)=17.48$, $p<.001$] and MODALITY [$F_{SHORT}(1,18)=188.68$, $p<.001$; $F_{LONG}(1,18)=165.61$, $p<.001$], and a significant EMOTION X MODALITY interaction [$F_{SHORT}(3,54)=8.57$, $p=.001$; $F_{LONG}(3,54)= 12.79$, $p<.001$].

Follow-up tests performed on the interaction established that for declarative intonation, happy prosody was produced significantly higher in F_0 than neutral, sad, and angry prosody, and angry prosody was significantly higher than neutral and sad prosody. For interrogative intonation, fewer distinctions were present as a function of mean F_0, although sad prosody consistently exhibited a significantly lower F_0 than happy, angry, and neutral prosody (which did not differ significantly). Neutral prosody was also significantly lower in F_0 than happy and angry prosody for interrogatives. Interrogatives were always shown to be significantly higher in mean F_0 than declaratives for all emotion types.

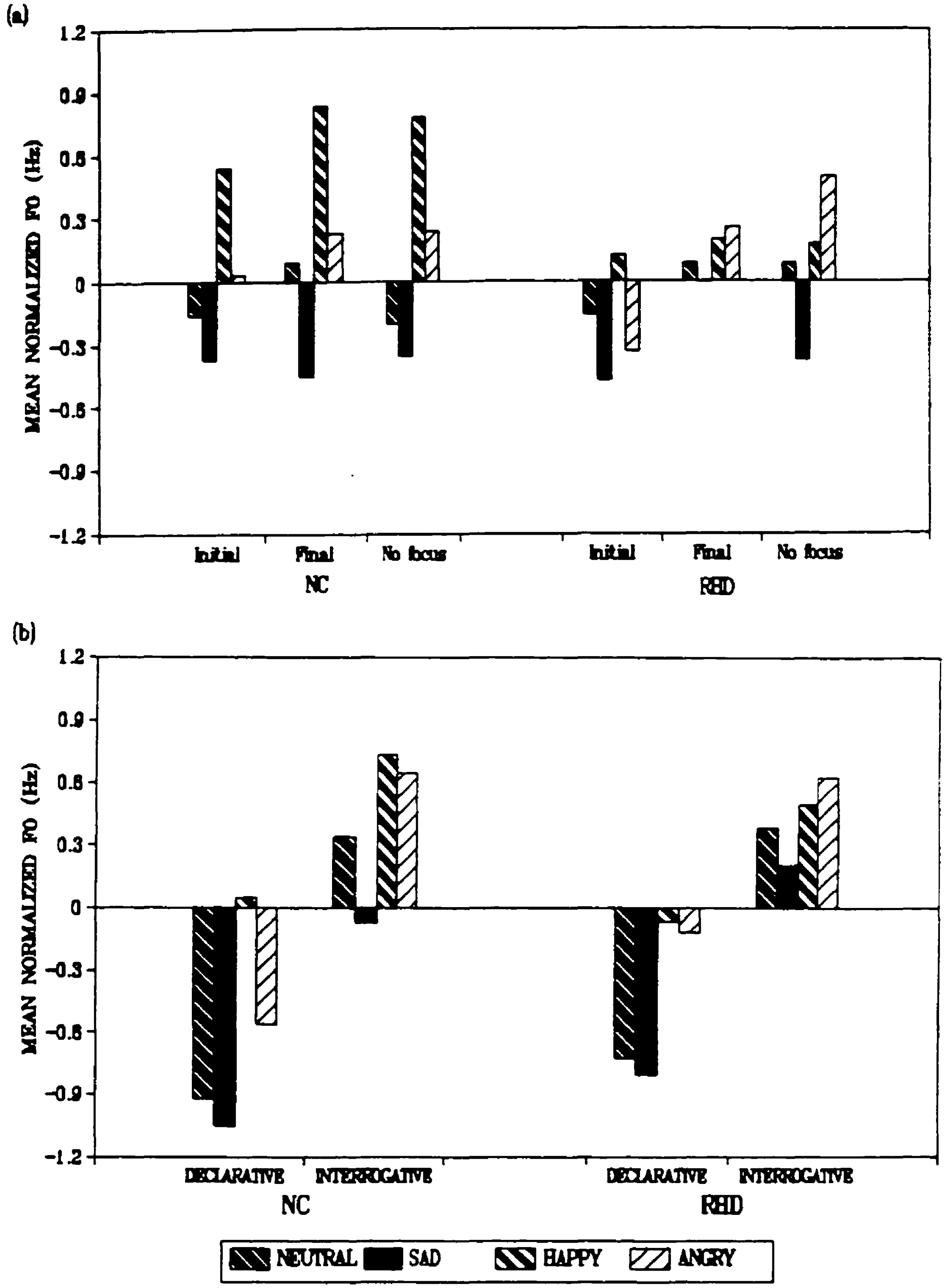
A significant MODALITY X FOCUS interaction was also found for both the short and long stimuli [$F_{SHORT}(2,36)=45.47$, $p<.001$; $F_{LONG}(2,36)= 20.49$, $p<.001$]. This interaction, for declaratives, was explained by a significantly lower F_0 in utterances containing sentence-initial focus relative to the other two focus versions (sentences with sentence-initial focus displayed uniformly low F_0 values on all keywords following the focused item). For short stimuli, utterances without focus were also significantly lower in F_0 than utterances with sentence-final focus. In comparison, interrogative sentences with sentence-initial focus were always significantly *higher* in mean F_0 than sentences with terminal focus; (for long stimuli, sentences with initial focus were also higher than sentences without focus). Generally, these results reflect the diverging contours described for declaratives and interrogatives when sentence-initial focus is present, described earlier in the discussion of keyword F_0.

Several interactions reached statistical significance only for the production of short sentences. These were: GROUP X EMOTION [$F(3,54) = 3.19, p<.05$], GROUP X FOCUS [$F(2,36) = 3.75, p<.05$], EMOTION X FOCUS [$F(6,108) = 2.23, p<.05$], and GROUP X EMOTION X FOCUS [$F(6,108) = 2.79, p<.05$]. Closer inspection of the three-way interaction (displayed in Figure 18a) indicated that the RHD patients produced happy prosody with a significantly lower mean F_0 than the NC subjects for two of the three focus versions (sentence-final and no focus). No between-group differences were observed in the production of neutral, sad, or angry prosody as a function of emphasis placement.

The ability of each group to differentiate discrete emotions based on mean F_0 was additionally shown to vary: although normal subjects always expressed happy prosody with a significantly elevated mean F_0 relative to that of neutral and sad prosody, the RHD patients elevated the F_0 of happy sentences only when compared to sad sentences, and then only in one context (when sentence-initial focus was present)[5]. These findings suggest that the RHD patients may have been disturbed in setting F_0 at a level appropriate to happy prosody when producing

[5]The NC subjects also produced significantly higher mean F_0 values for the following emotions: 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 F_0 than sad prosody in one context (utterances without focus).

Figure 18. Interaction of (a) group, emotion, and focus for short stimuli, and (b) group, emotion, and sentence modality for long stimuli.



short sentences, leading to a reduction in the contrastiveness of the four emotions in terms of mean F_0.

Finally, analysis of mean F_0 for only the long stimuli revealed a significant interaction of GROUP X EMOTION X MODALITY [$F(3,54)=3.30$, $p<.05$], shown graphically in Figure 18b. Between-groups, this interaction was explained by the significantly higher F_0 at which angry prosody was produced by the RHD group relative to the NC group for declarative utterances. Perhaps more importantly, the capacity of each group to express F_0 distinctions among the four emotions varied considerably as a function of sentence modality for the long utterances. Although both groups were relatively successful at differentiating the four emotions when produced in conjunction with declarative intonation (happy and angry prosody were both significantly higher in F_0 than sad and neutral prosody for both groups, and happy was higher than angry for the NC group), almost no emotional distinctions were produced by the RHD group in conjunction with interrogative intonation. Only mean F_0 for angry prosody exceeded sad prosody for the RHD patients when interrogatives were produced, whereas the NC subjects distinguished significantly among three of the emotions (i.e., happy, angry, and neutral were elevated in F_0 relative to sad, and happy was elevated in F_0 relative to neutral).

(iii) F_0 Range

Normalized F_0 range values produced by each group for the short and long stimuli are supplied in Table 15, as a function of emotional mode, sentence modality, and focus location. Separate (GROUP X EMOTION X MODALITY X FOCUS) ANOVAs performed on the short and long data each produced significant main effects for GROUP [$F_{SHORT}(1,18)=5.33$, $p<.05$; $F_{LONG}(1,18)=5.00$, $p<.05$] and EMOTION [$F_{SHORT}(3,54)=17.14$, $p<.001$; $F_{LONG}(3,54) = 18.02$, $p<.001$]. The group main effect was explained by the significantly smaller F_0 range produced by the RHD patients across stimulus types relative to the NC subjects. The emotion factor interacted with other variables in both the short and long analyses and is described below.

Both analyses additionally yielded a significant MODALITY X FOCUS interaction [$F_{SHORT}(2,36)=7.80$, $p<.001$; $F_{LONG}(2,36)= 11.46$, $p=.001$]. For both the short and long sentences, F_0 range was shown to be significantly larger in declaratives with initial focus relative to those without focus. By comparison, F_0 range was significantly *smaller* in interrogatives with initial focus relative to those with final focus. These data suggest that emphasis produced at the beginning of utterances exerts a substantial influence on the amount of F_0 modulation permitted after the focused item.

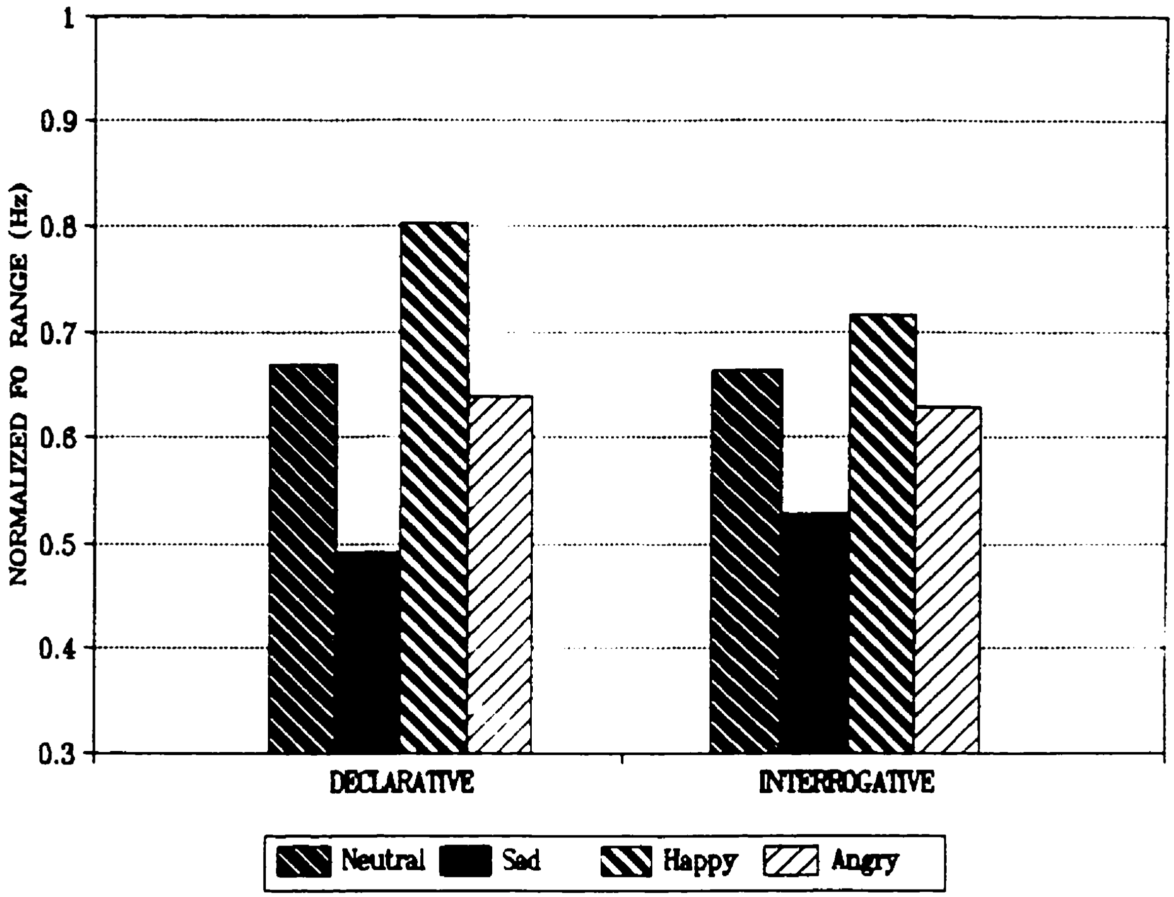
Table 15. Mean F_0 range of short and long utterances produced by the NC and RHD groups as a function of emotion type, sentence modality, and focus location.

			NC				RHD			
Length	*Mod*	*Focus*	N	S	H	A	N	S	H	A
Short	Declar	Initial	0.82	0.51	0.98	0.73	0.68	0.49	0.69	0.58
		Final	0.84	0.53	0.95	0.75	0.65	0.49	0.69	0.63
		None	0.53	0.55	0.83	0.65	0.51	0.38	0.68	0.50
	Inter	Initial	0.63	0.58	0.77	0.58	0.56	0.43	0.64	0.46
		Final	0.75	0.66	0.91	0.82	0.64	0.45	0.61	0.58
		None	0.81	0.61	0.80	0.74	0.58	0.44	0.56	0.61
Long	Declar	Initial	0.87	0.51	1.01	0.73	0.51	0.48	0.60	0.60
		Final	0.61	0.53	0.77	0.66	0.56	0.51	0.62	0.59
		None	0.48	0.43	0.83	0.70	0.50	0.42	0.59	0.56
	Inter	Initial	0.61	0.58	0.77	0.64	0.56	0.47	0.57	0.48
		Final	0.81	0.59	0.91	0.79	0.65	0.42	0.64	0.64
		None	0.75	0.65	0.81	0.84	0.57	0.52	0.65	0.61

Note: Declar= Declarative, Inter=Interrogative, N=Neutral, S=Sad, H=Happy, A=Angry. Range values were normalized for inter-subject differences in mean F_0 by dividing the F_0 range of each utterance by its mean F_0, for each speaker.

For only the short stimuli, a significant main effect of FOCUS [$F(2,36)=5.00$, $p=.01$](described above within the interaction of focus and modality) and a significant interaction of EMOTION X MODALITY [$F(3,54)=3.45$, $p<.05$] were noted. Post hoc exploration of the interaction (presented in Figure 19) indicated that for both declaratives and interrogatives, sad prosody was distinct in its significantly restricted F_0 range relative to the other

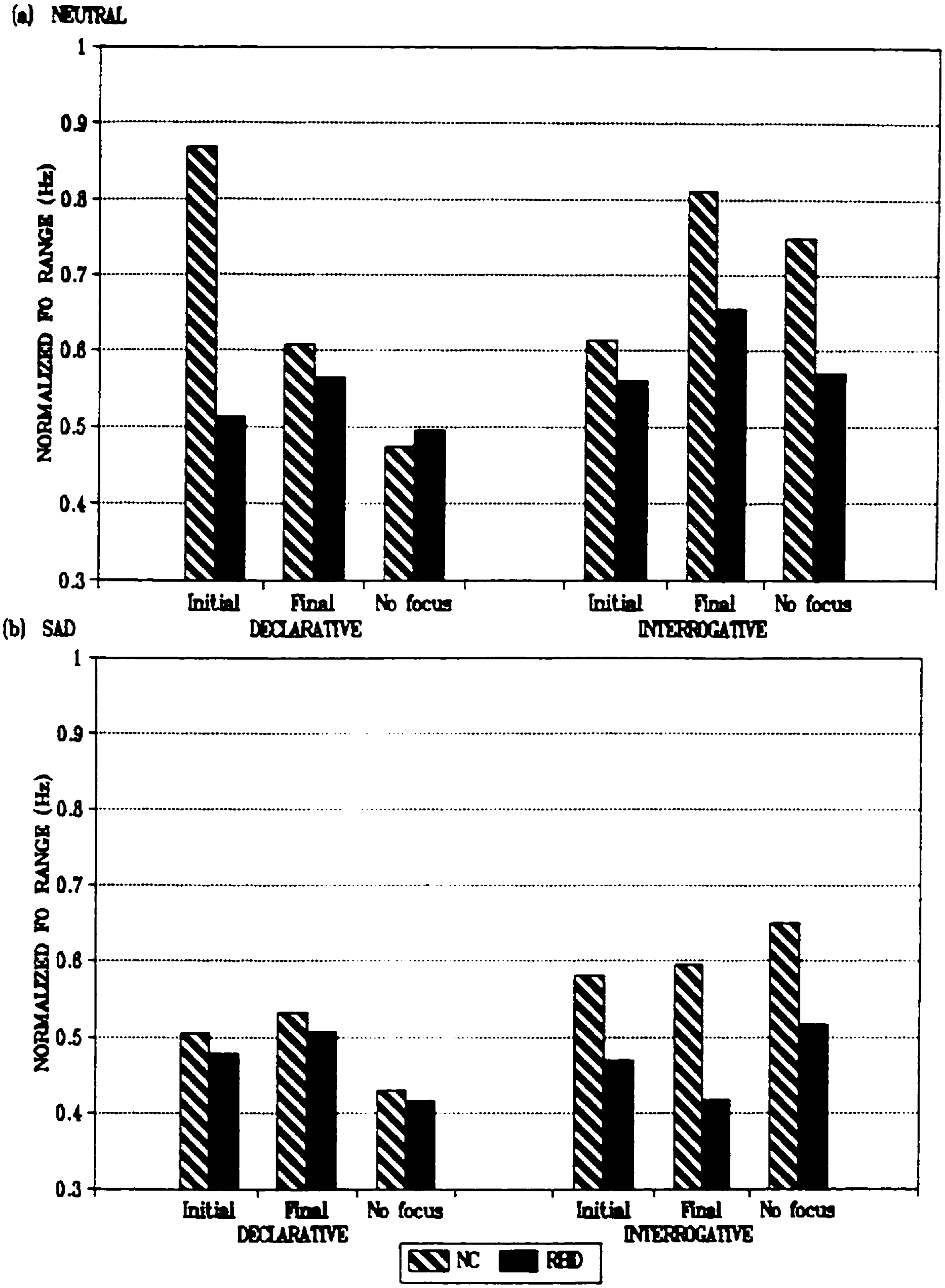
Figure 19. Interaction of emotion and sentence modality for short stimuli.

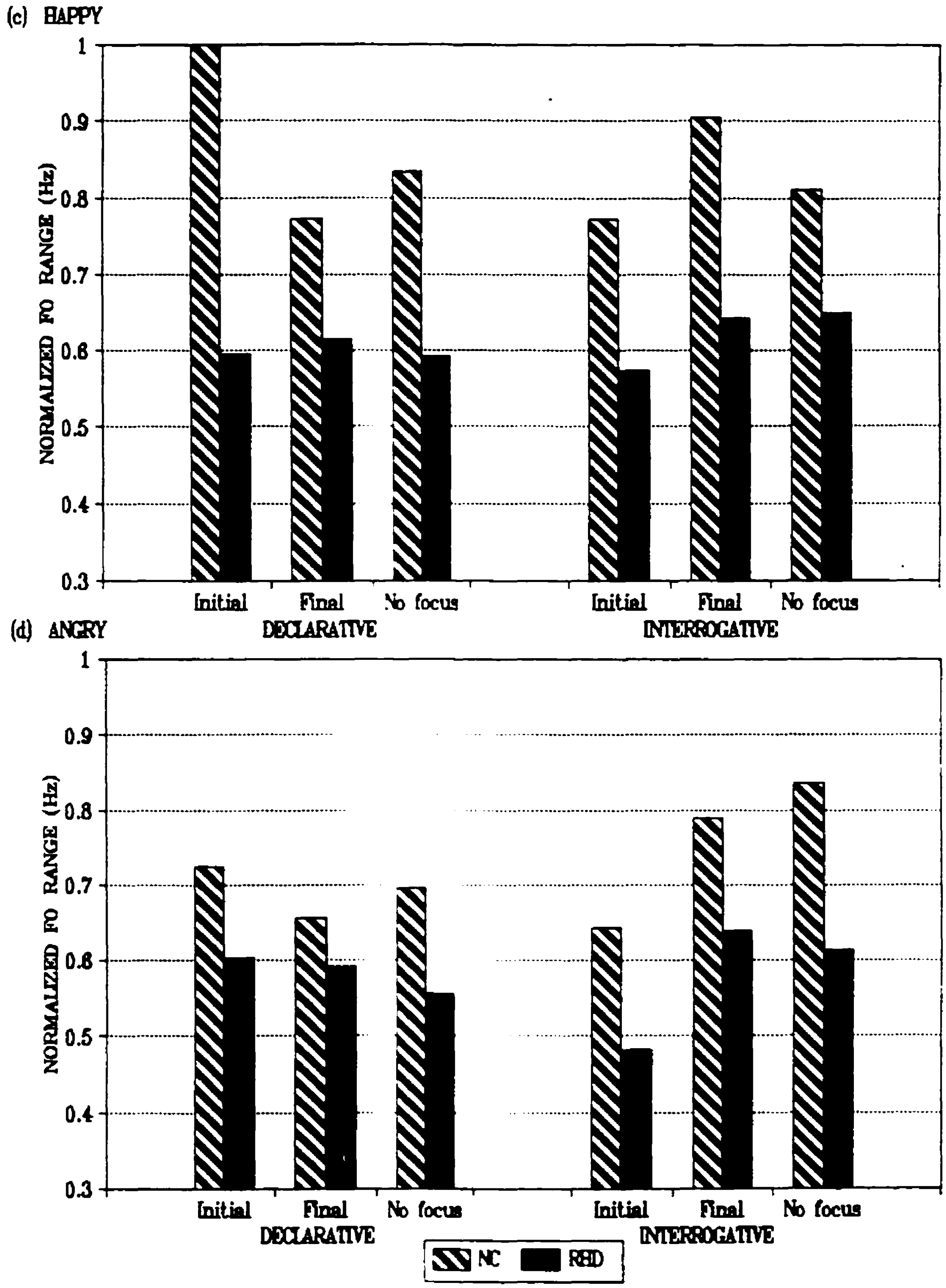


three emotions. Happy prosody also exhibited a significantly larger range relative to angry prosody for both modality types, and relative to neutral prosody for declaratives. Comparison of range values in matched declarative and interrogative sentences suggested that only happy prosody differed as a function of sentence modality, range values proving significantly constrained when spoken as an interrogative as compared to a declarative (see Figure 19). No interactions with the group factor were noted for the production of the short stimuli.

Finally, for only the long stimuli, the following significant effects were observed: EMOTION X FOCUS [$F(6,108)=2.80$, $p<.05$], EMOTION X MODALITY X FOCUS [$F(6,108)=3.83$, $p<.01$], GROUP X MODALITY X FOCUS [$F(2,36)=3.84$, $p<.05$], and GROUP X EMOTION X MODALITY X FOCUS [$F(6,108)=2.82$, $p<.05$]. Inspection of the four-way interaction (presented in Figure 20 (a-d), by emotion) revealed that the RHD patients produced significantly restricted F_0 ranges for four of the six versions of happy prosody (declaratives with initial and no focus, interrogatives with initial and final focus--see Figure 20c) when compared to the normal subjects. As well, the RHD group produced less F_0 variation than the NC group in one version of neutral prosody (declaratives with initial emphasis--Figure 20a) and one version of angry prosody (interrogatives without focus--Figure 20d). The RHD patients did not surpass the NC subjects with respect to F_0 range for any comparison in these data. These findings complement those reported for F_0 accent and mean F_0 measures, demonstrating that RHD speakers exhibited a select deficit in modulating the F_0 correlates of happy prosody relative to normal speakers.

Figure 20. Interaction of group, emotion, sentence modality, and focus for long stimuli spoken in a (a) neutral, (b) sad, (c) happy, and (d) neutral context.




(c) HAPPY
NORMALIZED F0 RANGE (Hz)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Initial
Final
No focus
DECLARATIVE
Initial
Final
No focus
INTERROGATIVE
(d) ANGRY
NORMALIZED F0 RANGE (Hz)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Initial
Final
No focus
DECLARATIVE
Initial
Final
No focus
INTERROGATIVE
NC
RHD

DISCUSSION

Disparate conclusions in the literature about the right hemisphere's role in encoding speech prosody may reflect, in part, myriad experimental approaches undertaken to investigate this question. When viewed across studies, differences in the length, complexity, and functional attributes of prosodic patterns elicited from RHD patients have led to conflicting views about the right hemisphere's role in prosodic processing (e.g., Behrens, 1989; Ross, 1981; Shapiro & Danly, 1985; Van Lancker & Sidtis, 1992). This uncertainty is compounded by the suggestion that different structural properties of tone and non-tone languages may further determine right hemisphere participation on prosodic tasks (Ross et al., 1988; Gandour et al., 1995). Delineating the capacity of RHD native English speakers to modulate prosody when multiple distinctions--both linguistic and emotional, and those encompassing both "local" and "global" acoustic domains--are present concurrently 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).

Emphatic stress cues, although manifest locally at the word level, are encoded and perceived as a relative measure of acoustic change occurring throughout a sentential unit (e.g., Bolinger, 1961; 1978). Typically, alterations in both the duration and F_0 characteristics of focused syllables contribute to emphatic highlighting in speech (Behrens, 1988; Cooper et al., 1985). The success of RHD patients in using these parameters to signal focus was investigated by computing duration and F_0 measures at key sites located throughout the tokens elicited in Experiment 1. In this manner, both the relative magnitude and direction of acoustic change implemented by the RHD and NC subjects could be monitored at strategic points in their utterances.

As noted on frequent occasions in the present results, the use of both duration and F_0 cues in emphasis was highly dependent on differences in focus position, sentence modality, and utterance length (discussed below). Despite this natural variability, the production of emphatic stress by the RHD patients patterned in a relatively similar manner to that of the normal subjects for both duration and F_0 in many important respects. For example, emphasized vowels in both sentence-initial and sentence-final words tended to be longer in duration than matched vowels without emphasis for both groups (Behrens, 1988; Cooper et al., 1985; Eady & Cooper, 1986; Ouellette & Baum, 1993). As well, vowels located in "intervening" words were always reduced in duration relative to those in sentence-initial and sentence-final positions for both the RHD and NC groups.

The use of mean F_0 in signaling emphasis also proved comparable for the NC and RHD subjects in many respects; for example, both groups produced a large post-focus drop in F_0 subsequent to sentence-initial words in declarative utterances (Cooper et al., 1985; Eady & Cooper, 1986). Thus, despite certain irregularities in their speech (discussed below), the RHD patients' qualitatively normal use of duration and F_0 to signal focus across a wide variety of stimulus types (a corpus of nearly 1000 utterances was scrutinized for each group) suggests that knowledge of emphatic stress production was largely spared in the current RHD sample (Behrens, 1988; Ouellette & Baum, 1993). This result is interpreted as evidence of minimal right hemisphere involvement in the processes that encode linguistic stress features at the word level, as hypothesized previously (Behrens, 1988, 1989; Emmorey, 1987; Van Lancker, 1980).

Not all research is harmonious with the conclusion that emphatic stress production is largely spared in patients with unilateral right-hemisphere dysfunction. However, those investigations reporting disturbed production of emphatic stress in RHD patients are conspicuous in their analysis of relatively few experimental stimuli, and their reliance on auditory-perceptual ratings (rather than acoustic analysis) to determine the presence of a prosodic deficit (Brådvik et al., 1991; Bryan, 1989; Weintraub et al., 1981). As suggested by Behrens' (1988) data, the production of "normal" acoustic cues to emphasis by RHD patients may not always be reflected in emphasis cues that are of normal *perceptual* saliency to the listener; this discrepancy between the physical stimulus and its auditory percept may have led to divergent claims in the literature from studies employing acoustic versus perceptual evaluation procedures. Indeed, the current finding that F_0 accents on focused items, although appropriate in general shape and direction between the two groups, were significantly reduced in overall *magnitude* when produced by the RHD patients (group main effect) demonstrates that emphatic stress cues were not *entirely* normal in the clinical sample. These differences, if subjected to *perceptual* evaluation, could well prove compatible with the listener's impression of reduced or impaired emphasis production by the RHD speakers, as suggested previously (Brådvik et al., 1991; Bryan, 1989; Weintraub et al., 1981).

Further abnormalities in the production of emphatic stress cues by the RHD patients were occasionally noted. In general terms, these irregularities were characterized by inconsistent use of specific acoustic features across stimulus types, and reduced acoustic distinctiveness within the stimulus sets produced by the RHD patients. For example, RHD patients failed at times to lengthen initial focused vowels or raise their F_0 to the extent displayed by normal subjects (see Ouellette & Baum, 1993, for similar trends in the use of duration). Furthermore, utterances varying in focus location tended to be less distinct with respect to duration and F_0 at keyword positions where emphatic contrasts are typically most salient (i.e., sentence-initial and sentence-final for the current study) for the RHD speakers relative to normal speakers. Curiously, this trend was especially prevalent when interrogatives were elicited from the RHD patients, in which case differences in focus distribution were often completely unmarked by the two acoustic parameters at designated positions within the stimuli.

Thus, despite resembling normal speakers in the general manner in which duration and F_0 cues were manipulated to mark linguistic focus, RHD patients tended to exploit *fewer* normal cues to emphasis in their productions (Behrens, 1988; Hird & Kirsner, 1993). This pattern points to quantitative differences between the RHD and NC groups in implementing duration and F_0 to signal linguistic focus, differences potentially exacerbated by the (assumed) increased processing demands of encoding interrogative intonation simultaneously. Again, it is unclear whether these acoustic irregularities would reduce the normalcy of the RHD patients' emphatic stress tokens when examined from the *listener's* perspective (see Behrens, 1988, for a discussion of this point). Work currently underway seeks to correlate the acoustic data reported herein with auditory-perceptual evaluations of the RHD patients' speech.

In addition to emphatic stress, analysis of keyword measures provided data on the RHD patients' ability to produce sentence modality distinctions, that is, to contrast declarative and interrogative utterances acoustically. Examination of the normative data reported herein indicates that these two illocutionary modes were distinguished reliably by the *direction* of F_0 fluctuation observed in the final 150 milliseconds of the utterance, interrogative contours exhibiting a marked rise in F_0 versus a terminal decline for declarative contours (Behrens, 1989; Eady & Cooper, 1986; Ohala, 1984; Studdert-Kennedy & Hadding, 1973). However, sentence modality markers were also prevalent at points preceding the terminal; most notably, for sentences with sentence-initial focus, the direction of the post-focus F_0 excursion corresponded to the marked (interrogative) or unmarked (declarative) direction of the F_0 terminal (Eady & Cooper, 1986). When emphasis did not occur sentence-initially, acoustic comparisons at specific points of declarative and interrogative contours were less instructive; however, a tendency to produce "non-falling" contours for interrogative utterances was generally observed (Behrens, 1989). This trend frequently rendered interrogatives significantly higher in F_0 than declaratives in the latter half of the utterance (Hadding-Koch & Studdert-Kennedy, 1964; Majewski, 1968). These data reaffirm that, despite the recognized importance of terminal F_0 in signaling modality distinctions, the acoustic concomitants of modality specifications are encoded very early on in sentence production, and are highly inter-dependent on other prosodic features in the signal such as emphatic stress (Eady & Cooper, 1986; McRoberts et al., 1995).

In keeping with the emphatic stress data, declarative and interrogative contours elicited from the RHD patients described herein were generally reflective of normal patterns of cue usage, with some exceptions. Although the RHD patients demonstrated consistent use of F_0 in differentiating declaratives and interrogatives at the utterance terminal (interrogatives were always significantly higher in F_0 than declaratives), their ability to encode interrogative contours in conjunction with emotional prosody appeared to be aberrant at times when compared to the normal subjects. For example, RHD patients marked interrogation with a significantly smaller terminal F_0 rise than NC subjects when speaking in a sad tone, and failed altogether to produce a terminal rise in F_0 when speaking in a happy tone. In contrast, measures of terminal *declination* underlying declarative utterances did not significantly differ for the two groups as a function of emotional content.

Thus, although the RHD patients were successful at marking sentence modality distinctions in a vast number of utterances through terminal F_0 features (thereby demonstrating intact knowledge of linguistic categorical features of intonation contours), the concurrent demands of modulating F_0 to signal interrogation (the "marked" intonational feature according to Lieberman, 1967) and emotional attributes highly reliant on F_0 parameters may have proven difficult for the RHD patients, yielding abnormal prosodic attributes in their speech. Given that happy and sad prosody are normally expressed with maximal and minimal F_0 modulation (respectively) within the current stimulus set, it is perhaps not surprising that the production of these two emotional tones interfered with F_0 signaling for other (linguistic) purposes by the RHD patients. Thus, the prosodic "load" specified by individual acoustic features such as mean F_0 may be a factor in the expressive performance of RHD patients. The suggestion that linguistic and emotional sentence prosody exert an *interactive* influence on RHD speakers represents the first instance of such evidence for native speakers of a non-tone language, although several researchers have commented on the impact of phonemic tone on affective signaling in RHD speakers of tone languages (Edmondson et al., 1987; Gandour et al., 1995).

Based on prior research (Fairbanks & Pronovost, 1939; Scherer, 1986; Williams & Stevens, 1972), the ability of the RHD patients to produce emotional-prosodic contrasts was investigated through long-term (i.e., "global") measures of acoustic change averaged across the utterance as a whole. Generally, it was shown that the four affect types elicited in the present study separated well as a function of speech rate, mean F_0, and F_0 range. When compared to neutral prosody (i.e., the "unmarked" affect within the present stimulus set), happy prosody tended to be spoken at a relatively similar (high) rate but with an elevated mean F_0 and larger F_0 dispersion overall. Sad prosody was always produced much slower in rate and with a diminished F_0 range relative to neutral prosody, although mean F_0 tended to be relatively similar between these two forms. Finally, angry prosody resembled neutral prosody in F_0 range, but was spoken at a significantly reduced rate and with a significantly elevated mean F_0. Thus, when compared to emotionally-inflected utterances, neutral utterances were characterized by a relatively high speech rate, low mean F_0, and an intermediate amount of F_0 variation. This acoustic profile conforms generally to previous descriptions in the literature on emotional communication[6] (Fairbanks & Hoaglin, 1941; Fairbanks & Pronovost, 1939; Frick, 1985; Pell & Baum, 1997b; Scherer,1986; Williams & Stevens, 1972).

The RHD subjects' proficiency at modulating the acoustic concomitants of discrete emotions was of prime interest in the present study, in light of multiple claims that this skill is impaired subsequent to right hemisphere dysfunction (Borod et al., 1990; Edmondson et al., 1987; Gandour et al., 1995; Ross, 1981;

---

[6]As described for emphatic stress, it is noteworthy that the number of acoustic distinctions observed among the four emotions was highly constrained by the shape of the intonation contour, far fewer emotional distinctions occurring in interrogative contours than in declarative contours. These data further illustrate how the F_0 features of linguistic and affective specifications interact in the speech signal.

Ross et al., 1988; Shapiro & Danly, 1985). Examination of the speech rate and mean F_0 data indicated that *general* dimensions of the emotional utterances produced by the RHD subjects were in keeping with those produced by the normal subjects (i.e., there was no group main effect for either measure)(Baum & Pell, 1997; Gandour et al., 1995). Interestingly, however, the RHD patients' use of both rate and mean F_0 cues departed from normalcy for specific emotion types. In the case of speech rate, RHD adults produced sad prosody (the emotion most distinctive in its slow rate relative to neutral prosody) significantly faster than normal subjects. These rate abnormalities resulted in poor separation of the four emotions as a function of speech rate for the RHD subjects when compared to the normal subjects.

In the case of mean F_0, RHD patients demonstrated abnormal production of emotional tones that required *elevating* F_0 in relation to the baseline established by neutral prosody (i.e., the "unmarked" affect. Gandour et al., 1995). More precisely, RHD patients failed to raise F_0 to the level of normal subjects when signaling happy prosody in short sentences, and elevated F_0 significantly further than normal subjects when signaling angry prosody in long sentences. Collectively, these findings highlight subtle difficulties in the ability of RHD patients to regulate timing and mean F_0 to demarcate discrete emotional tones in a manner paralleling normal speakers (Colsher et al., 1987; Dykstra et al., 1995; Edmondson et al., 1987; Shapiro & Danly, 1985). Perhaps more importantly, these findings suggest that for the RHD patients, aberrant production of emotional contours emerged predominantly for those forms representing *extreme* departures from unmarked (i.e., neutral) contours in normal speech for each individual acoustic parameter (i.e., speech rate for sad prosody, mean F_0 for happy and angry prosody; see below for a more detailed discusssion of this point).

The RHD patients' capacity to modulate F_0 *variation* throughout an utterance (measured here by means of F_0 range) demonstrated a more pervasive impairment than that observed for speech rate or mean F_0. Analysis of F_0 range measures across the full range of stimuli generated in Experiment 1 indicated that the RHD patients displayed significantly restricted intonational variation in their productions when compared to normal utterances (i.e., a group main effect emerged). 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; Shapiro & Danly, 1985).

However, in addition to a generalized restriction of F_0 range, the RHD patients reported herein displayed an emotion-specific impairment in modulating F_0 range. The production of happy prosody, an emotion characterized by extensive F_0 variation relative to the other three emotions, was often significantly reduced in F_0 range in the RHD patients' productions when compared to matched tokens produced by the normal subjects. Thus, the F_0 range data supply further support for the contention that acoustic features representing extreme departures from unmarked (i.e., non-affective) contours in emotional speech pose particular difficulties for RHD patients, a pattern suggested in the speech rate and mean F_0 data.

The idea that intonation contours largely devoid of emotional overtones constitute the "unmarked" affective mode in speech, and that increases or decreases within various acoustic dimensions (such as mean F_0, F_0 range, tempo) serve to inflect neutral contours with emotional information, is not new (Bolinger, 1986; Gandour et al., 1995; Ladd. 1980). Evidence that RHD patients retained the capacity to implement local (emphatic stress) and global (sentence modality) *categorical* changes in linguistic contours, and produced emotional contours that patterned in a relatively similar manner to those of normal subjects when viewed generally, speaks to the integrity of prosodic representations for encoding purposes in the current RHD sample.

Rather, abnormalities in the RHD patients' speech appeared to stem from deficient control of *continuous* aspects of prosodic phenomena (Blonder et al., 1995); as discussed, F_0 variation, a gradient feature, was restricted in a generalized manner as determined both locally (on emphasized keywords) and globally (the utterance as a whole).[7] Moreover, long-term acoustic parameters (speech rate, mean F_0, F_0 range) were aberrant for the RHD speakers predominantly for those stimuli displaying extreme inflections from the "neutral" position along a

---

[7]Although emphatic stress may be linguistically-assigned and therefore categorical in nature, the observation that increased acoustic change associated with emphasis leads to the perception of greater emphasis (e.g., Bolinger. 1961) suggests that continuous acoustic properties also contribute to emphatic stress.

continuum of change specified for each individual cue. This apparent inability to distinguish prosodic categories normally in terms of distinct landmarks within the relevant acoustic dimensions culminated in fewer inter-stimulus differences for the RHD patients overall, or a tendency for utterances varying in emotional or emphatic content to "converge" acoustically. This apparent implementation deficit, if evaluated by means of auditory-perceptual analyses, may lead to subjective impressions of reduced intonation or "flat affect" in the present RHD speakers (Borod et al., 1990; Ross, 1981; Ross & Mesulam, 1979).

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). Similarly, it is not unexpected that few studies have reported 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 (Behrens, 1989; Blonder et al., 1995; Cooper et al., 1984; Gandour et al., 1995; Hird & Kirsner, 1993; Lebrun et al., 1985; Ryalls et al., 1987; but cf. Brådvik et al., 1991; Shapiro & Danly, 1985; Weintraub et al., 1981). This pattern of performance is predicted by the relative importance of gradient acoustic phenomena in signaling affective versus "grammatical" prosodic meanings in verbal behaviour. Remaining conflicts in the literature on linguistic intonation may reflect, in part, idiosyncracies in cue use within experimentally-defined speaking modes; in addition, the likelihood that "neutral" utterances are seldom completely devoid of speaker affect (and thus, emotional highlighting via continuous acoustic information) may have further contributed to divergent conclusions in this literature. As demonstrated in the present investigation, the very presence of emphatic information in non-affective utterances may be sufficient to evoke expressive dysprosodies in RHD patients, underscoring that it is the nature of the cues and not their linguistic or affective function that is *most* central to the disturbance experienced by these patients.

The prospect that RHD patients experience difficulties modulating gradient aspects of prosodic patterns is also compatible with the hypothesis that the *domain* of prosody expression acts as a determinant of prosody disruption in RHD speakers (Baum & Pell, 1997; Behrens, 1989; Emmorey, 1987; Gandour et al., 1992). Difficulties in regulating long-term properties of intonation contours exhibited by the RHD patients in this and prior studies contrast with the uniform observation that RHD patients retain the ability to produce phonemically-assigned stress and tone at the *syllabic* level (Behrens, 1988; Edmondson et al., 1987; Emmorey, 1987; Gandour et al., 1992; Hird & Kirsner, 1993; Lebrun et al., 1985; Ouellette & Baum, 1993; Ryalls & Reinvang, 1986). In light of the right hemisphere's hypothesized superiority for continuous aspects of vocal cues, these differences may reflect the increased prevalence of continuous acoustic information in relatively large prosodic units (i.e., intonational attributes) versus small prosodic units (i.e., categorical stress features), rather than a right hemisphere sensitivity to the size of the prosodic unit itself.[8] Such an interpretation renders the present comparison of "short" and "long" utterances less instructive for the RHD speakers, as both forms are realized by substantial modulation of continuous acoustic features in production (but see the results for interesting differences in the normative data related to sentence length). Indeed, a systematic link between sentence length and disruptions of prosody in the RHD patients did not generally emerge from these data, although previous reports suggest that such an association may be true for LHD aphasic patients, if tested (Danly et al., 1983; Gandour et al., 1989).

Finally, the hypothesis that duration and F_0 parameters of prosodic stimuli are independently lateralized to the left and right hemispheres of the brain (Robin et al., 1990; Van Lancker & Sidtis, 1992) does not obtain strong support from the present data (see also Pell & Baum, 1997b). As discussed, RHD subjects presented with deficits for both temporal and spectral aspects of prosody when used continuously, indicating that more than F_0 properties of intonation are disrupted subsequent to right hemisphere insult (Blonder at al., 1995; Dykstra et al., 1995). However, given the paramount importance of F_0 cues in signaling prosodic meaning in speech production (e.g., Bolinger, 1955; 1958; Lieberman,

---

[8]The "ambiguous" nature of emphatic stress, as a form composed of both categorical and continuous features and one assigned locally but encoded in relation to other constituents within the utterance, may require only partial involvement of the right hemisphere based on the present hypothesis.

1960), it is likely that the type of quantitative impairment attributed to RHD patients in the present research emerges more frequently for F_0 parameters, as suggested by the "cue hypothesis" (Robin et al., 1990; Van Lancker & Sidtis, 1992). Certainly, the idea that unilateral right-brain-damage yields a sensitivity to select elements of the physical form of prosodic contours is corroborated by the current results.

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 processing gradient 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. Research currently underway may help elucidate this issue.

EXPERIMENT 2

To test the ability of RHD subjects to *identify* the communicative import of linguistic and emotional prosody, six receptive tasks were presented in Experiment 2. Baseline tasks examined the ability of RHD patients, relative to NC and LHD subjects, to recognize linguistic and emotional stimuli as a function of concurrent alterations in prosodic content (related to focus, linguistic modality, and emotion, as in Experiment 1). Two additional tasks manipulated select acoustic properties of the same stimuli (F_0 or duration) to investigate the perceptual influence of each type of acoustic cue in identifying linguistic and emotional prosody relative to the baseline task.

METHODS

Subjects

Nine (9) RHD and 10 NC subjects who participated in Experiment 1 also took part in Experiment 2 (one RHD patient, R4, was unavailable to complete Experiment 2). In addition, a comparison group of eleven (11) unilaterally left-hemisphere-damaged subjects (mean=65.5 years, range=43-81) volunteered to participate in Experiment 2; the mean age of the three diagnostic groups (NC=66.1, RHD=64.7, LHD=65.5) did not significantly differ based on cursory inspection of the data. All LHD patients were right-handed English speakers who had undergone rehabilitation for aphasic deficits (7 nonfluent, 4 fluent) but who demonstrated good auditory comprehension of language, as determined by a subsection of the Psycholinguistic Assessment of Language (Caplan, 1992) and an aphasia screening. Table 16 summarizes basic clinical and demographic characteristics of the LHD group.

Clinical attributes of the LHD patients were determined in a similar manner to those described for the RHD patients. CT information confirmed the site and etiology of lesion within the left hemisphere. A puretone air conduction screening ensured that all subjects displayed acceptable hearing acuity. Only one LHD patient (L3) showed evidence of a contralateral visual neglect following administration of the Bells Test (Gauthier et al., 1989). Additional screening procedures included tasks of both spoken and written (single) word recognition, as well as auditory digit span. Again, only patients with relatively stable clinical features (i.e., at least 3 months post-CVA) were included in the study (mean=39 months post-onset, range=6-91).

The inclusion of a left-hemisphere-damaged control group in Experiment 2 permitted a better characterization of the deficits related specifically to RHD (and not just brain damage in general), allowing broader generalizations to be made about the perceptual lateralization of prosodic stimuli. The absence of a LHD control group in Experiment 1 was justified by the limitations of this group in encoding complex linguistic material, a position that has been adopted by other investigators in related studies (Gandour et al., 1995).

Table 16. Basic demographic and clinical characteristics of the LHD subjects participating in Experiment 2.

Sbj.	Sex	Age	Post-Onset	Lesion Site	Major Clinical Signs
L1	M	48	91	(L) parietal	(R) hemiparesis, mild-mod. nonfluent aphasia
L2	F	43	48	(L) frontoparietal with subcortical extension	(R) hemiparesis, severe nonfluent aphasia, severe verbal and oral apraxia
L3	F	63	23	(L) fronto-temporo-parietal	(R) hemiparesis, mod.-severe nonfluent aphasia, (R) visual neglect
L4	M	68	61	(L) parietal	(R) hemiplegia, severe nonfluent aphasia
L5	F	79	9	(L) frontoparietal	(R) hemiparesis, mod.-severe nonfluent aphasia, flat affect
L6	F	63	6	(L) frontoparietal	(R) hemiparesis, severe nonfluent aphasia
L7	F	67	29	(L) parietal	(R) hemiparesis, nonfluent aphasia
L8	F	81	46	(L) para-ventricular deep parietal	(R) hemiparesis, mild fluent aphasia
L9	F	79	31	(L) MCA	(R) hemiparesis, moderate fluent aphasia
L10	F	53	34	(L) basal ganglia	(R) hemiparesis, mild fluent (anomic) aphasia
L11	F	77	56	(L) MCA	(R) hemiparesis, mild fluent (anomic) aphasia

Note: age=years, post-onset=months.

Stimuli

Five sentences, each six syllables in length and ressembling those listed as "short" utterances in Experiment 1 (e.g., *Robert read the letter*) served as base stimuli in Experiment 2. To test the interaction of focus distribution (first, last, none), linguistic modality, (declarative, interrogative), and emotion (angry, sad, happy, neutral) on the perception of prosody, multiple renditions of each of the five stimuli were elicited by a normal speaker; 24 unique versions of each stimulus exhausted the potential combinations of these three factors.

An adult female speaker recorded the 120 utterances (5 X 24) onto digital audio tape in a soundproof chamber. The speaker was encouraged to produce each token several times to allow the experimenter an adequate sample from which to choose exemplars believed to be "on-target" for each of the three dimensions. The 120 utterances selected were then sampled and digitized onto disk using BLISS (20 kHz sampling rate, 9 kHz low-pass filter, 12-bit quantization).

Perceptual Rating Procedure

To ensure that the perceptual stimuli successfully conveyed the prescribed combinations of prosodic attributes, the 120 stimuli were rated independently by five phonetically-trained listeners using a checklist. Stimuli were randomized for order of presentation and then presented to raters over headphones by a computer. Each trial was played two consecutive times (separated by a 3.5 second interstimulus interval) to allow raters adequate opportunity to judge each token on the three dimensions of importance: location of emphasis within the sentence (first, last, none), sentence "type" (statement, question), and emotion (happy, sad, angry, neutral).

Inclusion criteria were set at four out of five correct responses for each of the three dimensions, per stimulus. Productions that did not meet these criteria upon initial presentation were re-recorded by the female speaker and judged by the same five raters on a subsequent occasion (at least one week later). To ensure that the entire set of 120 stimuli adequately projected the desired prosodic meanings, three additional rating sessions were required. Upon completion of the fourth session, 117 (98%) of the stimuli were judged "correct" on all three prosodic dimensions by the five raters, whereas 3 of the stimuli (2%) met criterion on two of the dimensions but failed to meet criterion on the third. However, as the three "problematic" stimuli did not show a bias towards failure on a specific prosodic category, they were deemed acceptable for inclusion in the experiment.

Acoustic Analyses

The 120 perceptually-validated utterances were subjected to acoustic analyses to determine some of the physical differences underlying this set of stimuli. More importantly, however, an acoustic characterization of the perceptual stimuli was prerequisite to manipulating specific parameters of the stimuli important in the decoding of linguistic and emotional prosody (outlined in

Experimental Tasks). To illustrate both temporal and spectral changes in each prosodic contour, measures of duration and mean F_0 were extracted on the stressed vowel of each content word in a manner identical to that performed in Experiment 1. Visual inspection of the speech waveform enabled the duration of the target vowel (steady state and transition) to be determined by placing cursors at zero crossings at the vowel's onset and offset. Mean F_0 was calculated from the average of five contiguous glottal pulses isolated at the centre of each vowel. The F_0 of the utterance terminal (measured in the final 150 ms) and the total utterance duration were further noted. An acoustic profile of the F_0 and duration distinctions present in the perceptual stimuli is supplied in Tables 17 and 18, respectively. As may be seen, these data generally conform to normal patterns of cue usage in speech production as reported in detail in Experiment 1.

Experimental Tasks

The perception of prosodic cues by RHD, LHD, and NC subjects was assessed in six independent tasks: three tasks explored the subjects' ability to process local prosodic markers of emphatic stress (*Focus Perception*); and three tasks examined the perception of emotional vocal content by the same listeners (*Emotion Perception*). For both focus and emotion perception, subjects were presented a baseline task in which the stimuli were unaltered (i.e., those stimuli rated by the phonetically-trained listeners) and two tasks in which either the F_0 or duration parameters of the stimuli had been systematically modified (Ladd et al., 1985; Lieberman & Michaels, 1962). The precise manner in which the linguistic and emotional stimuli were acoustically manipulated differed somewhat (as described below), as did the type of judgement required of listeners on linguistic as opposed to emotional tasks. Otherwise, focus perception and emotion perception tasks were highly comparable.

Table 17. Mean F_0 of each stressed vowel and utterance terminal for the 120 perceptual stimuli, as a function of emotion, focus location, and sentence modality (in Hz, collapsed across the 5 items).

		V1 e.g. ***Ro**-bert*		**V2** *read*		**V3** *let-*		**Terminal** *ter*	
Emotion	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Neutral	None	225	229	201	221	180	187	162	320
	Initial	250	195	170	338	165	354	171	431
	Final	214	233	193	216	207	196	162	432
Sad	None	201	213	179	213	180	195	183	265
	Initial	190	196	171	271	171	269	174	297
	Final	218	209	189	195	182	198	172	282
Happy	None	439	322	187	260	409	289	194	533
	Initial	482	265	208	497	185	504	172	596
	Final	425	314	255	287	476	343	229	611
Angry	None	302	253	226	236	234	231	178	320
	Initial	309	303	202	338	193	348	176	411
	Final	267	284	217	257	244	284	177	488

Note: V=vowel, (.)=declarative utterances, (?)=interrogative utterances.

Table 18. Mean duration of each stressed vowel and total utterance duration for the 120 perceptual stimuli, as a function of emotion, focus location, and sentence modality (in milliseconds, collapsed across the 5 items).

		V1 *e.g. Ro-bert*		**V2** ***read***		**V3** *let-ter*		**Utterance Total**	
Emotion	*Focus*	(.)	(?)	(.)	(?)	(.)	(?)	(.)	(?)
Neutral	None	103	106	86	69	110	125	1169	1144
	Initial	132	132	71	81	87	108	1138	1190
	Final	93	93	66	68	122	127	1110	1180
Sad	None	126	167	78	79	143	141	1311	1487
	Initial	190	170	78	90	96	114	1340	1349
	Final	118	112	81	73	148	137	1408	1349
Happy	None	99	111	83	75	124	105	1102	1157
	Initial	150	149	93	75	120	113	1258	1217
	Final	95	96	79	74	146	121	1253	1224
Angry	None	113	125	63	62	132	115	1050	1141
	Initial	165	143	75	72	111	121	1205	1235
	Final	91	92	63	56	122	124	1103	1172

Note: V=vowel, (.)=declarative utterances, (?)=interrogative utterances.

1) Focus Perception

The ability of subjects to recognize the position of contrastive stress features within a sentence was tested in three independent tasks as a function of linguistic modality (declarative, interrogative) and emotional tone (angry, sad, happy, neutral). Only utterances conveying sentence-initial and sentence-final focus were presented in linguistic tasks: combined exhaustively with the other two factors (modality, emotion), this resulted in a total of 80 trials per task (5 items X 2 focus locations X 2 sentence modalities X 4 emotions). Sentences rated as lacking focus were not presented in experimental tasks, but were essential in implementing acoustic modifications to the stimuli presented in the "D-neutral" and "F-neutral" tasks (described below). Although experimental stimuli always exhibited either sentence-initial or sentence-final focus (as confirmed by the phonetically-trained raters), subjects were offered three choices from which to form a response, including a "no focus" category. The following three tasks were presented to each subject:

(i) Baseline - This task assessed the ability of brain-injured subjects to process emphatic stress cues from stimuli in which all potential prosodic features (but no semantic features) contributed to emphasis in the speech signal. This task was designed to elucidate whether the perception of emphasis is relatively more susceptible to focal right- or left-hemisphere insult, and reveal whether the cues to discrete emotions or sentence modality distinctions perceptually influence those that convey sentence focus (McRoberts et al., 1995).

(ii) Neutral Duration ("D-Neutral") - The contribution of changes in vowel duration to the perception of focus was tested using stimuli in which the durational correlates to focus were acoustically "neutralized" at each focus location prior to presentation. Based on the acoustic data derived for the 120 stimuli (see

Table 18), the vowel durations of utterances *with* focus were adjusted to correspond to those of their *unfocused* counterparts (i.e., the same vowel in the same position of "no focus" utterances, all other factors remaining constant). Thus, each utterance lacking focus served as the point of comparison for two utterances with focus (one sentence-initial, one sentence-final). Prior to this comparison, vowel durations were corrected for differences in speaking rate between stimuli by expressing each as a proportion of the total utterance duration. Pitch periods were then either removed from the centre of the focused vowel or added to it to reflect the duration of the vowel spoken in the unfocused context. Cuts were made at zero-crossings to ensure that there was no audible evidence of the editing. This process resulted in a set of 80 stimuli that varied naturally in F_0 (and other) cues to focus, but for which durational distinctions were neutralized with respect to the location of emphasis within the utterance.

(iii) Neutral F_0 (F-Neutral) - The significance of F_0 cues in identifying focus location was explored by again manipulating the acoustic form of the base stimuli to effectively "neutralize" the effect of F_0 cues prior to perceptual identification. In preparing the stimuli, the F_0 of each utterance was extracted via an autocorrelation algorithm and the LPC spectrum for the utterance was computed automatically. A vocoding software programme was then utilized to replace the F_0 contour of the focused syllable with that of its unfocused homologue in the same keyword position, and the utterance was re-synthesized. In this manner, stimuli presented in the neutral-F_0 condition retained the temporal characteristics to focus across the utterance, but F_0 cues were rendered neutral with respect to the focus location.

2) Emotion Perception

To establish whether brain-damaged subjects were capable of interpreting the emotional intent of speech, particularly as a function of differences in focus location (initial, final, none) and sentence modality (declarative, interrogative), three tasks of 90 trials each were presented (5 items X 3 focus locations X 2 sentence modalities X 3 emotions). Only stimuli judged to be angry, sad, and happy by the phonetically-trained raters were presented in emotional tasks: stimuli designated by the phonetically-trained individuals as conveying "no emotion" were again used as landmarks for implementing various acoustic manipulations to the experimental stimuli. However, as was the case in the focus perception tasks, a "neutral" (i.e., no emotion) category was offered to subjects as an additional response option in judging the stimuli. The following three tasks were presented to each subject:

(i) Baseline - An initial task established each diagnostic group's ability to interpret the emotional intent of semantically-neutral sentences in which *all* potential prosodic cues signaled the emotional meaning. The potential effects of focus location and linguistic modality on the identification of emotional material was additionally explored for each diagnostic group.

(ii) Neutral Duration ("D-Neutral) - As changes *throughout* an utterance contribute to emotional signals, the effect of temporal changes in perceiving affective content was studied by "neutralizing" the vowel length of *all* content words in emotional utterances. Based on the acoustic data collected for the 120 test stimuli (presented in Table 18), the duration of each stressed vowel in emotional utterances was compared to the duration of the same vowel in its "nonemotional" homologue (i.e., each neutral utterance was compared to three emotional utterances, all other factors remaining constant). The vowel durations of emotional utterances were then adjusted (where necessary) to reflect the values of the same vowels spoken in a nonemotional context. Again, this was achieved by adding or removing pitch periods at the centre of the vowel. This process culminated in a set of stimuli varying naturally in F_0 cues to the target emotion, but for which duration cues were no longer indicative of the target emotion.

(iii) Neutral F_0 ("F-Neutral") - To investigate the effects of pitch change on the perceptual recognition of emotion, emotional stimuli were manipulated in an attempt to "neutralize" the contribution of pitch cues while preserving other distinctions in the stimuli. To achieve this, the F_0 of each utterance was extracted via an autocorrelation algorithm and the LPC spectrum for the utterance was computed. The utterances were then vocoded by replacing the F_0 contour of each utterance (representing the topline, or peak F_0 values of all content words in the utterance) with that of its non-affective homologue (all other factors remaining constant). The utterance was then re-synthesized, maintaining the temporal pattern of the signal but eradicating the effect of pitch cues in projecting the speaker's affect. This task tested the importance of the F_0 contour in conveying vocal emotion, and evaluated the extent to which discrete emotions are identifiable from their temporal characteristics.

Procedure

Subjects were tested individually on two separate occasions, three of the perceptual tasks being presented during each visit. The order in which focus and emotion perception tasks were presented was randomized for each subject and counterbalanced among the three groups, with the restriction that at least one linguistic or emotional task be presented during each session.

For all six tasks, subjects listened to auditory stimuli one at a time (separated by a five second interval) over headphones. Auditory stimuli were fully randomized within each task and presented by a computer, which recorded the accuracy of each subject's responses. Linguistic or emotional judgements were indicated by pushing a button on a response board (aligned vertically). Buttons were labelled both verbally (focus perception: first, last, none; emotion perception: angry, sad, happy, neutral) and with a corresponding pictogram (e.g., a facial expression). Five practice trials ensured that subjects were comfortable with the task demands and were oriented to the positioning of response buttons. A five minute break was set at the half-way point of each task.

RESULTS

Accuracy data emanating from the three tasks of focus perception and the three tasks of emotion perception were examined independently using analysis of variance (ANOVA) techniques. Post hoc analyses, where appropriate, were performed using Tukey's method ($p<.05$).[9]

1) Focus Perception

The accuracy of the three subject groups in locating emphatic stress cues as a function of concurrent alterations in prosodic content (i.e., sentence modality, emotion) was explored independently for each of the three focus perception tasks (Baseline, D-Neutral, F-Neutral). Data from each task were examined using a 3 X 2 X 2 X 4 repeated measures ANOVA in a full factorial design: for each analysis, GROUP (NC, RHD, LHD) constituted the between-subjects factor, and FOCUS (initial, final), MODALITY (declarative, interrogative) and EMOTION (angry, sad, happy, neutral) served as the within-subjects factors. A subsequent 3 X 2 X 3 ANOVA explored how the acoustic manipulation of the stimuli influenced each group's perception of focus irrespective of modality and emotion.

---

[9]Given the complexity of the experimental design, an exhaustive comparison of all cell means was not always insightful for certain high-order interactions. Accordingly, only those comparisons deemed most theoretically relevant are reported below.

(i) Baseline

The identification of emphatic stress location in the Baseline condition provided a measure of each group's ability to perceive and interpret the acoustic correlates to linguistic focus in natural (i.e., acoustically unaltered) stimuli. Table 19 supplies the mean percentage of correct responses observed for each group in identifying focus in sentence-initial and sentence-final position, as a function of sentence modality and emotion type.

The (GROUP X FOCUS X MODALITY X EMOTION) ANOVA performed on these data revealed a significant main effect for GROUP [$F(2,27) = 10.94$, $p<.001$]. Follow-up analyses revealed that both the NC and RHD subjects were significantly better than the LHD aphasic subjects at identifying the position of emphatic stress cues within simple (unaltered) utterances. The accuracy of the NC and RHD groups did not significantly differ on this task. No interactive effects with the GROUP factor were found in this analysis.

The analysis also yielded significant main effects for MODALITY [$F(1,27) = 12.83$, $p=.001$] and EMOTION [$F(3,81) = 27.1$, $p<.001$] and significant interactive effects of FOCUS X EMOTION [$F(3,81) = 10.4$, $p<.001$], FOCUS X MODALITY [$F(1,27) = 15.31$, $p=.001$], MODALITY X EMOTION [$F(3,81) = 5.64$, $p=.001$], and FOCUS X MODALITY X EMOTION [$F(3,81) = 9.61$, $p<.001$]. A graphic depiction of the three-way interaction is supplied in Figure 21.

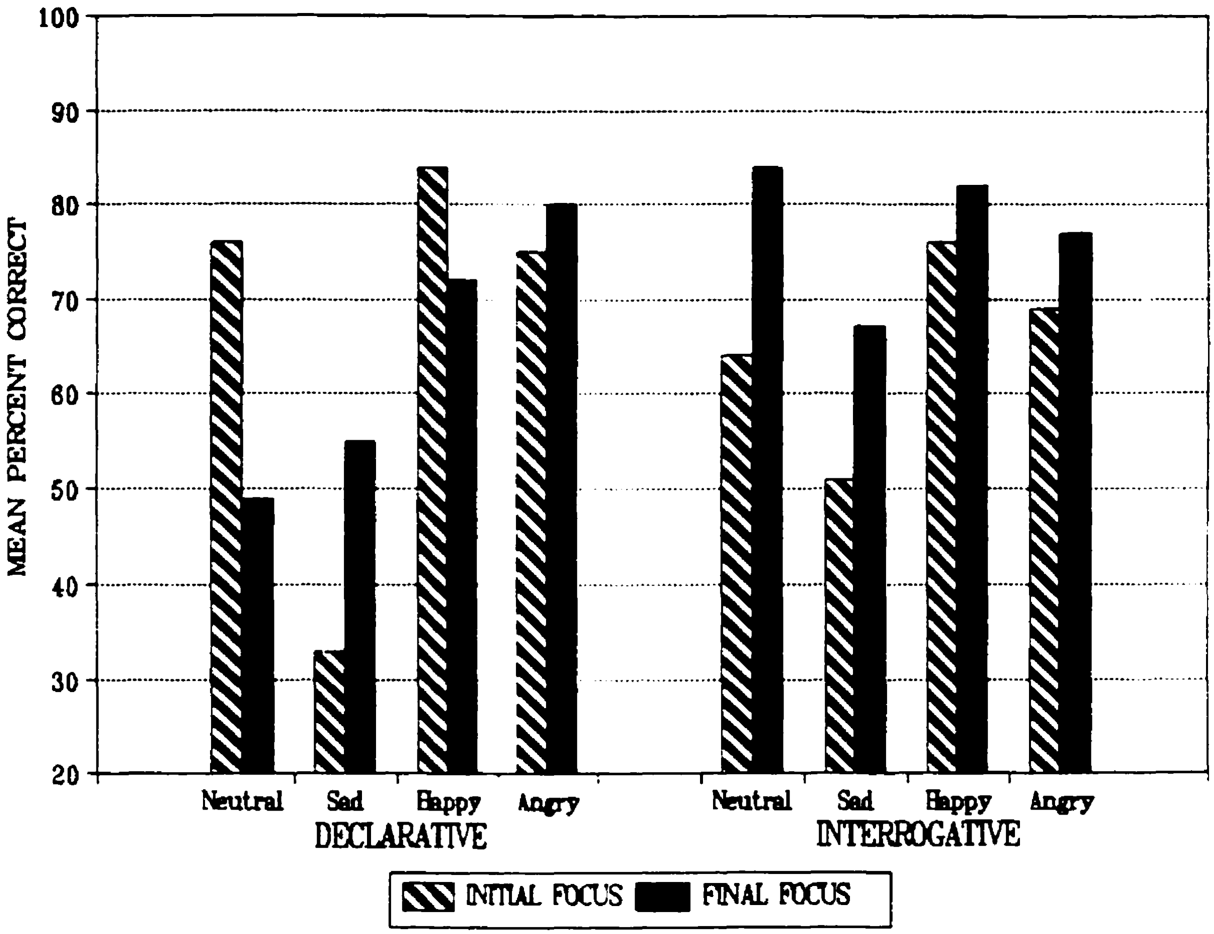
Table 19. Percent correct responses (by group) in identifying sentence-initial and sentence-final focus on the Baseline task, as a function of sentence modality and emotion type.

		PERCENT CORRECT					
		NC		RHD		LHD	
Modality	*Emotion*	*Initial*	*Final*	*Initial*	*Final*	*Initial*	*Final*
Declar.	Neutral	80	52	95	57	58	43*
	Sad	28*	68	58	55	18*	44*
	Happy	94	82	98	77	66	60
	Angry	76	90	87	89	69	65
Inter.	Neutral	66	92	78	91	51	74
	Sad	58	74	72	77	31*	53
	Happy	88	92	93	87	52	69
	Angry	62	88	93	87	58	61
MEAN (SD)		69 (21)	80 (14)	84 (14)	78 (14)	50 (17)	59 (11)

Note: (*) indicates performance expected by chance for a three alternative forced choice paradigm based on 95% confidence limits, where chance=33% ± 11 (binomial distribution).

Post hoc examination of the FOCUS X MODALITY X EMOTION interaction considered the influence of sentence modality and emotional distinctions on the subjects' ability to detect focus at each sentence position individually. For sentence-*initial* focus, recognition was significantly higher when presented in conjunction with happy, angry, and neutral prosody relative to sad prosody for declaratives. For interrogatives, sentence-initial focus was identified

Figure 21. Interaction of emotion, sentence modality, and focus on the perception of emphatic stress on the Baseline task.



significantly better in happy sentences than sad sentences. For sentence-*final* emphasis, accuracy was significantly higher for happy and angry stimuli relative to neutral stimuli, and for angry stimuli relative to sad stimuli, when declarative sentences were presented. However, when interrogative utterances were presented, the recognition of sentence-terminal emphasis did not vary as a function of emotion type, accuracy scores at this sentence position remaining relatively high for interrogatives across emotion types (see Figure 21).

(ii) Neutral Duration (D-Neutral)

It is generally accepted that increases in vowel length represent one of the primary correlates of linguistic focus (e.g., Eady & Cooper, 1986; Klatt, 1976). Accordingly, the D-Neutral task explored the extent to which subjects could determine focus position in utterances in which the temporal determinants of focus were absent (i.e., the duration of focused vowels was rendered congruent to that of unfocused items in matched utterances). Table 20 presents the accuracy performance of the three subject groups on the D-Neutral task, as a function of sentence modality and emotion type.

As exemplified by Table 20, the pattern of responses observed in the D-Neutral task closely resembled those described for the Baseline task. Statistical inspection of these data produced a significant main effect for GROUP [$F(2,27) = 4.42$, $p<.05$]. Post hoc examination of this effect revealed a significant decrement in the overall performance of the LHD group relative to that of both the NC and RHD groups, which did not significantly differ. A significant GROUP X FOCUS X EMOTION interaction was also found in this analysis [$F(6,81) = 3.32, p<.01$]; this interaction is presented in a graphic format in Figure 22.

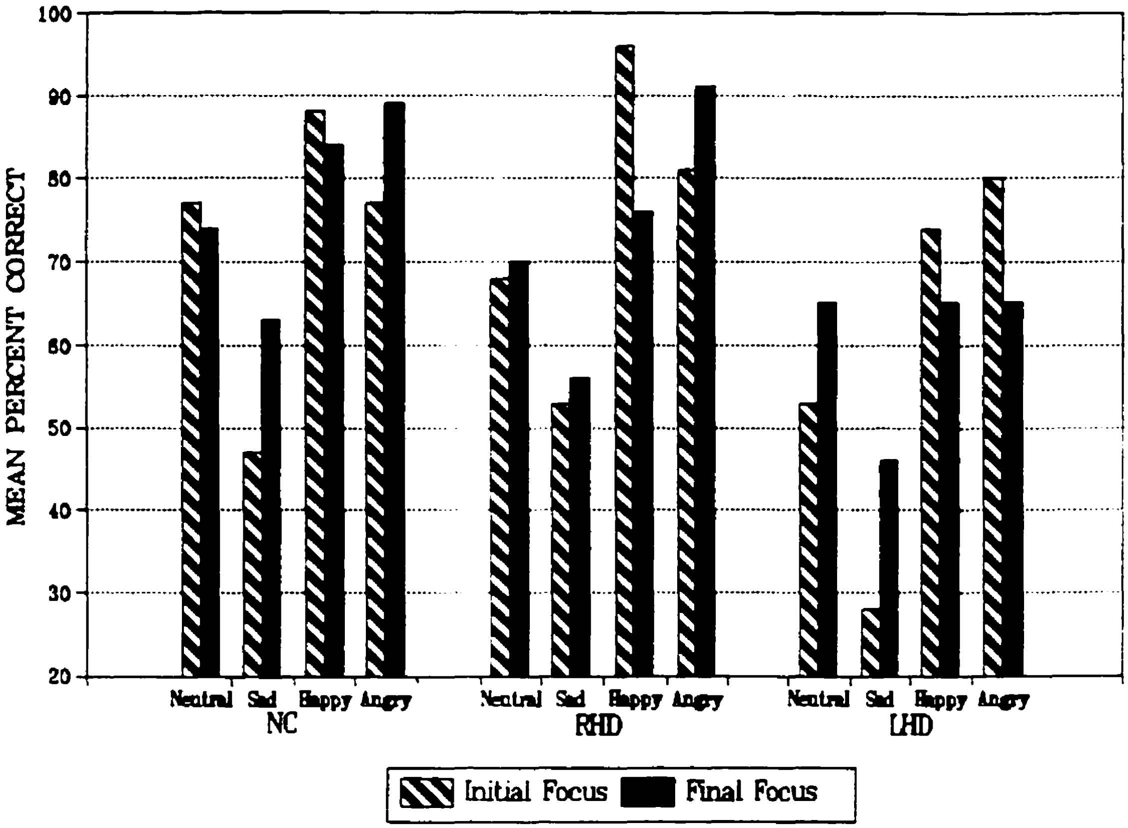
Table 20. Percent correct responses (by group) in identifying sentence-initial and sentence-final focus on the D-Neutral task, as a function of sentence modality and emotion type.

		PERCENT CORRECT					
		NC		RHD		LHD	
Modality	*Emotion*	*Initial*	*Final*	*Initial*	*Final*	*Initial*	*Final*
Declar.	Neutral	86	50	71	53	55	45
	Sad	24*	62	44*	56	24*	53
	Happy	96	76	98	69	76	60
	Angry	76	90	84	91	89	65
Inter.	Neutral	68	98	64	87	51	84
	Sad	70	64	62	56	33*	40*
	Happy	80	92	93	82	71	69
	Angry	78	88	78	91	71	65
MEAN (SD)		72 (21)	78 (17)	74 (18)	73 (17)	59 (22)	60 (14)

Note: (*) indicates performance expected by chance for a three alternative forced choice paradigm based on 95% confidence limits, where chance=33% ± 11 (binomial distribution).

Pairwise comparisons among the means examined how differences in focus location and emotional prosody affected the performance of each subject group.

Figure 22. Interaction of group, emotion, and focus on the perception of emphatic stress on the D-Neutral task.



For the NC group, accuracy in locating focus position was unaffected by the emotional mode of utterances on the D-Neutral task. However, both the RHD and LHD patients were less able to detect sentence-*initial* emphasis when sad prosody was presented than when happy or angry prosody was presented on the D-Neutral task, when normal temporal markers to prominence were unavailable; (LHD subjects further differed for sad and neutral prosody). The RHD patients were also impaired in the recognition of sentence-*final* emphasis for sad prosody relative to angry prosody, although the perceptual accuracy of the LHD group for sentence-final focus did not differ as a function of emotion or focus location on the D-Neutral task (see Figure 22).

Significant main effects for MODALITY [$F(1,27) = 7.23, p=.01$] and EMOTION [$F(3,81) = 45.28, p<.001$] were additionally observed, and several interactive effects proved significant: FOCUS X MODALITY [$F(1,27) = 13.02, p=.001$], FOCUS X EMOTION [$F(3,81) = 7.16, p<.001$], MODALITY X EMOTION [$F(3,81) = 7.87, p<.001$], and FOCUS X MODALITY X EMOTION [$F(3,81) = 27.94, p<.001$]. Exploration of the three-way interaction (reported in full in Appendix D) revealed a qualitatively similar pattern to that described on the Baseline task. Namely, the perception of emphasis tended to be most facilitated by happy and angry prosody, and least facilitated by sad and neutral prosody, with one exception; when subjects were required to identify sentence-final emphasis from interrogative stimuli, few differences were noted in subjects' accuracy across the four emotion types. This pattern suggests a link between the perceptual markers of sentence-final focus and interrogation which was not dependent upon duration parameters of the stimuli (which were neutralized on the D-Neutral task).

(iii) Neutral F_0 (F-Neutral)

Previous investigations suggest that a relatively large pitch excursion on (or around) linguistic material of communicative importance within a speech stream constitutes the primary perceptual determinant of sentential focus (e.g., Bolinger, 1958). Through the presentation of stimuli in which changes in pitch associated with focused items were effectively "neutralized" prior to the experiment, the F-Neutral task sought to assess the significance of these localized pitch cues on each group's perception of emphasis position. The mean accuracy of each group in identifying sentence-initial and sentence-final focus on the F-Neutral task is displayed in Table 21, as a function of sentence modality and emotion.

As may be seen, accuracy scores on the F-Neutral task were generally low overall; the absence of pitch excursions on focused items appeared to have a particularly detrimental effect on the performance of the NC group, resulting in a much smaller margin of accuracy differentiating the three groups. The ANOVA conducted on this data set uncovered a significant main effect for GROUP [$F(2,27) = 5.71, p<.01$]. *A posteriori* inspection of these data indicated that the RHD patients were significantly more accurate than the LHD patients at locating emphasis position within utterances which were devoid of the natural pitch excursions associated with prominence. However, neither the RHD nor the LHD groups differed significantly from the NC group overall, a pattern possibly reflecting the disproportionate difficulty experienced by the NC group on the F-Neutral task.

Table 21. Percent correct responses (by group) in identifying sentence-initial and sentence-final focus on the F-Neutral task, as a function of sentence modality and emotion type.

		PERCENT CORRECT					
		NC		RHD		LHD	
Modality	*Emotion*	*Initial*	*Final*	*Initial*	*Final*	*Initial*	*Final*
Declar.	Neutral	70	26*	76	40*	51	20*
	Sad	56	46	56	53	35*	36*
	Happy	56	68	71	80	53	56
	Angry	78	84	84	82	73	65
Inter.	Neutral	12*	88	29*	73	35*	75
	Sad	14*	54	40*	62	33*	44*
	Happy	62	90	69	78	56	65
	Angry	58	88	67	82	53	69
MEAN (SD)		51 (24)	68 (24)	62 (19)	69 (16)	49 (14)	54 (19)

Note: (*) indicates performance expected by chance for a three alternative forced choice paradigm based on 95% confidence limits, where chance=33% ± 11 (binomial distribution).

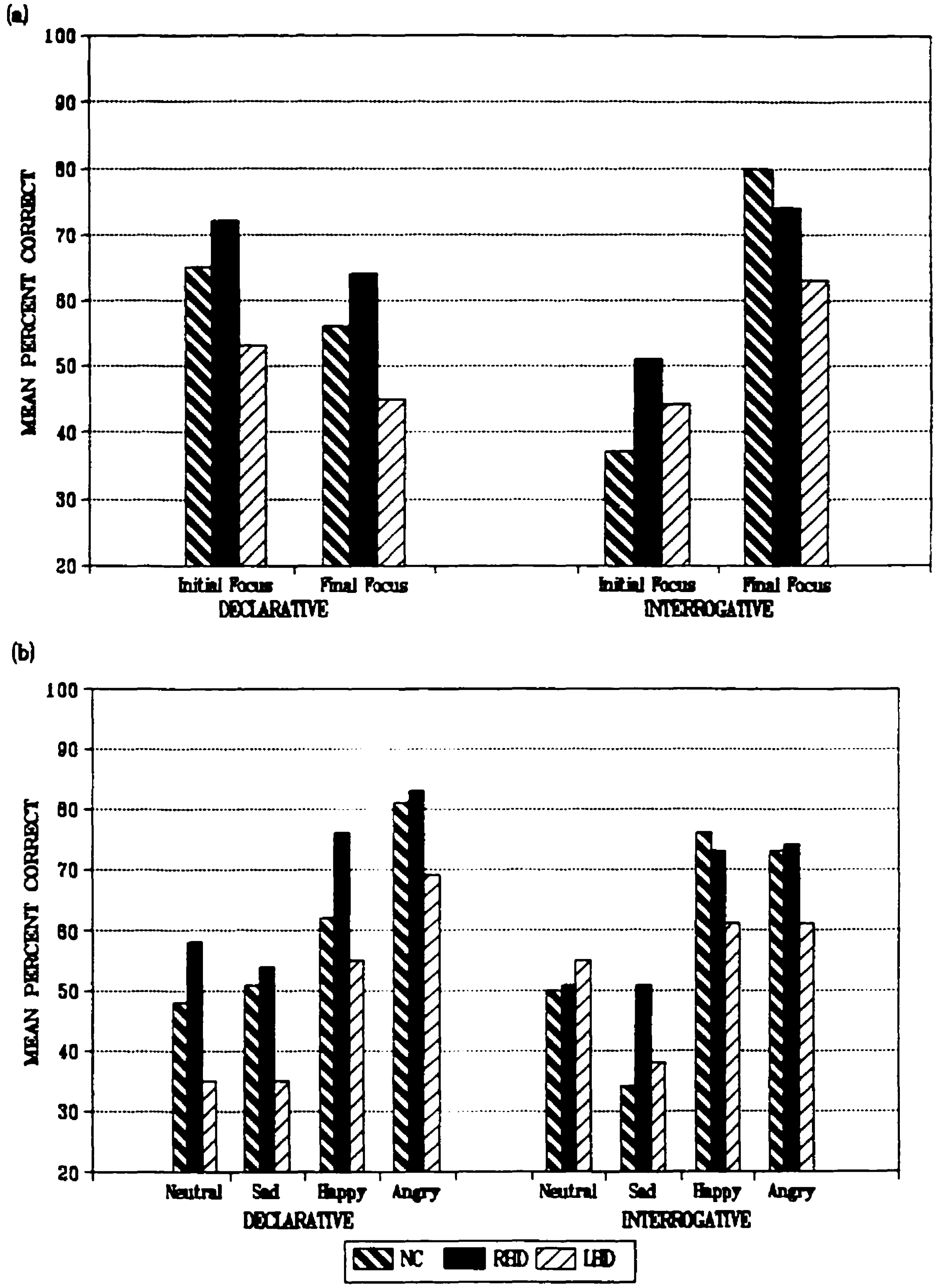
Significant interactions for GROUP X FOCUS X MODALITY [$F(2,27) = 3.39, p<.05$] and GROUP X MODALITY X EMOTION [$F(6,81) = 3.09, p<.01$]

were noted. For the GROUP X FOCUS X MODALITY interaction (displayed in Figure 23a), it was found that the RHD group was significantly better than the LHD group in detecting emphasis (both sentence-initial and final) when declarative intonation was presented, but no between-group differences in accuracy emerged when interrogative intonation was presented. Furthermore, all three groups identified sentence-final emphasis more reliably than sentence-initial emphasis when interrogatives were presented, whereas declarative intonation did not affect the accuracy of the three subjects groups in identifying focus position.

The GROUP X MODALITY X EMOTION interaction is displayed in Figure 23b. Follow-up tests indicated that for declaratives, the LHD group was significantly inferior in its recognition of sentence focus relative to the RHD group in three of the four emotional contexts (happy, sad, neutral). However, again no significant between-group differences in perceptual abilities were observed when focus identification took place in conjunction with interrogative intonation. Differences in sentence modality did not generally affect focus recognition for specific emotional modes for any group, although the LHD patients were significantly more accurate in locating focus in interrogatives than declaratives when neutral prosody was presented.

Significant main effects for FOCUS [$F(1,27) = 15.24, p=.001$] and EMOTION [$F(3,81) = 37.38, p<.001$] also emerged. The FOCUS main effect, which did not emerge from the analyses performed on the Baseline or D-Neutral data, was explained by the greater accuracy of subjects in determining focus

Figure 23. Interaction of (a) group, sentence modality, and focus and (b) group, emotion, and sentence modality on the perception of emphatic stress on the F-Neutral task.



location in sentence-final than in sentence-initial position; this finding suggests that pitch cues may be of relatively greater perceptual importance on items perceived sentence-initially. Other significant interactions produced by this analysis were: FOCUS X MODALITY [$F(1,27) = 71.58, p<.001$], MODALITY X EMOTION [$F(3,81) = 6.48, p=.001$], and FOCUS X MODALITY X EMOTION [$F(3,81) = 30.44, p<.001$].

Examination of the three-way interaction (reported in detail in Appendix D) revealed a relatively similar pattern to that described on the Baseline and D-Neutral tasks; recognition of emphasis was generally greatest for angry and happy stimuli and lowest for sad stimuli, for both declaratives and interrogatives. Exceptionally, happy prosody was not shown to facilitate recognition of sentence-initial focus in declaratives, a finding unique to the F-Neutral task in which stimuli lacked appropriate pitch accents on focused items. Finally, an association between interrogative intonation and superior recognition of sentence-final emphasis was again noted.

Summary

In summary, the recognition of emphatic stress cues by LHD aphasic patients was shown to be impaired relative to both matched RHD patients (3/3 tasks) and non-brain-damaged control subjects (2/3 tasks). On the F-Neutral task, the accuracy of the NC group approached that of the RHD and LHD groups and did not significantly differ from either group. In fact, overall accuracy measures for the RHD and NC groups did not differ significantly for any of the three tasks.

More generally, the perception of emphasis by both brain-damaged and non-brain-damaged subjects was shown to vary naturally as a function of the cues prevalent in different illocutionary or affective modes: focus recognition tended to be highest when happy and angry prosody were presented, and lowest when sad (and sometimes neutral) prosody were presented; and, interrogative intonation facilitated subjects' accuracy in detecting sentence-*final* focus relative to sentence-initial focus. The interaction of focus location, linguistic modality, and emotion type on the perception of emphasis was essentially similar for all three tasks, suggesting that the interplay of these various representations was largely unaffected by the acoustic manipulation of the stimuli. However, the emergence of group differences as a function of focus distribution, sentence modality and/or emotion type on the D-Neutral and F-Neutral tasks indicates that the effect of the acoustic manipulations on the three subjects groups was probably not entirely uniform (see discussion below).

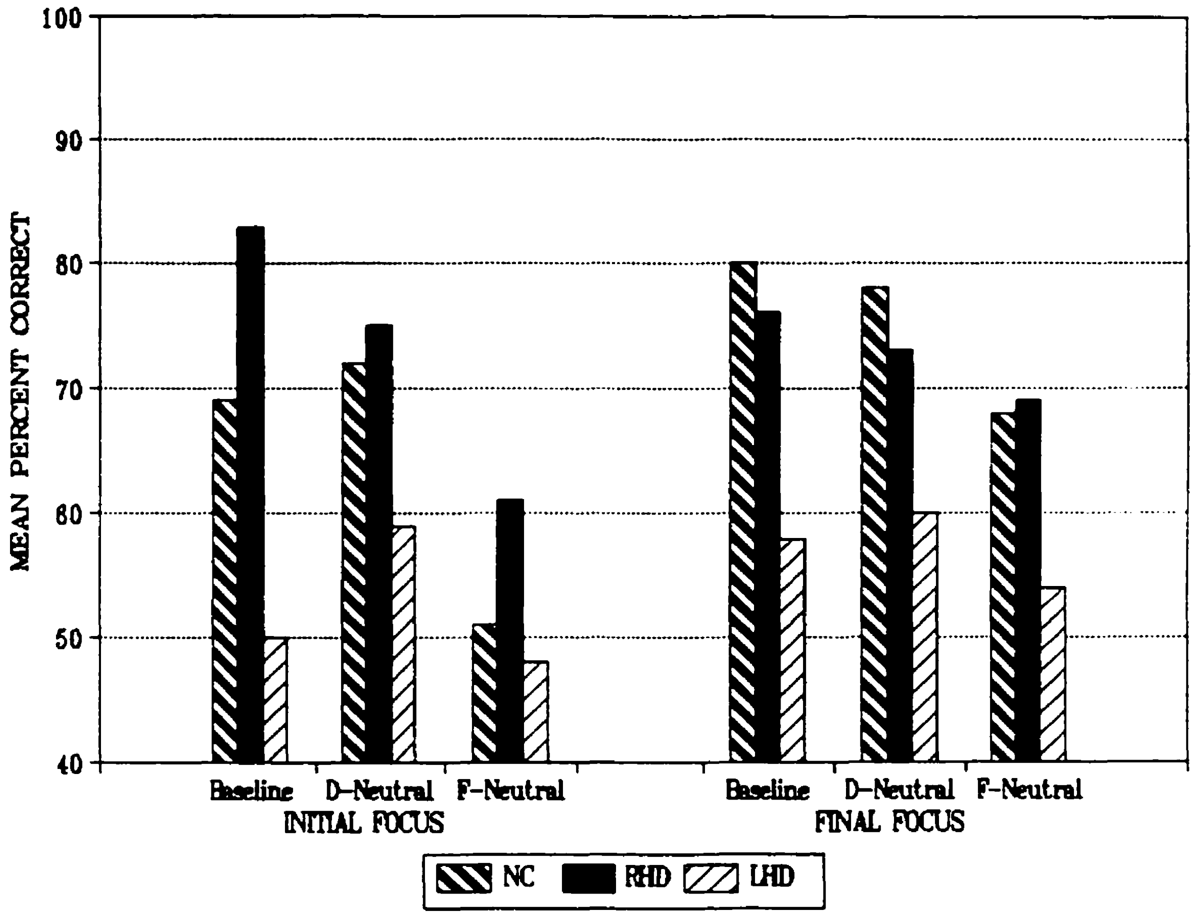
Effect of the Acoustic Manipulation of Emphasis Cues: TASK Factor

A direct test of the effects of manipulating individual acoustic cues to emphasis on the perception of focus location by brain-damaged and non-brain-damaged subjects was accomplished using a 3 X 2 X 3 ANOVA with repeated measures. Prior to this analysis, the data were collapsed within each task level across the two levels of "modality" and four levels of "emotion"; GROUP (NC, RHD, LHD) continued to serve as the between-subjects factor, and FOCUS (initial, final) and TASK (Baseline, D-Neutral, F-Neutral) served as within-subjects variables in this analysis. The accuracy performance of each group as a function of focus position and task level may be viewed in Figure 24.

As depicted in Figure 24, and as described previously for individual tasks, the ability of the LHD subjects to identify emphatic stress within simple utterances was inferior to that of the RHD and NC groups, irrespective of the acoustic composition of the stimuli (i.e., task level). Although minimal differences in the performance of each group were noted when the *duration* of focused items was neutralized (i.e., on the D-Neutral task relative to the Baseline), neutralizing the extent to which *pitch changes* normally occurred on focused items (i.e., F-Neutral task) led to a decrement in the performance of the three groups when compared to Baseline performance. This difference was especially pronounced for the NC and (to a lesser extent) RHD groups, but relatively small for the LHD group due to their already poor performance on the Baseline task.

The ANOVA performed on these data revealed a main effect for GROUP [$F(2,27) = 10.00$, $p=.001$]; this effect was explained by the impairment of the LHD group, overall, relative to both the RHD and NC groups (which did not significantly differ). Significant main effects for FOCUS [$F(1,27) = 4.62$, $p<.05$] and TASK [$F(2,54) = 15.10$, $p<.001$] also emerged. Post hoc inspection of these main effects suggested that sentence-final focus was identified significantly better

Figure 24. Overall accuracy of the NC, RHD, and LHD subjects in the Focus Perception condition as a function of focus position and task level.



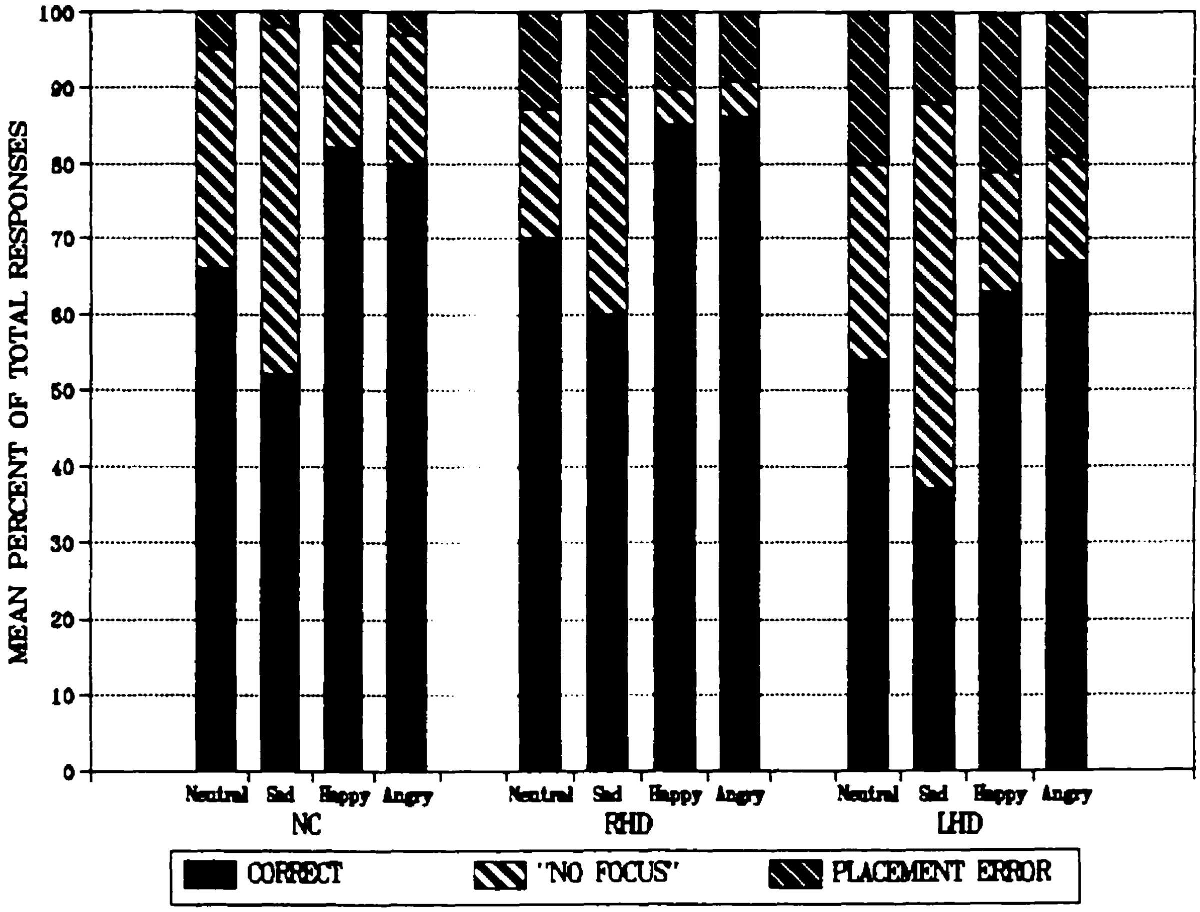
than sentence-initial focus, and that accuracy in both the Baseline and D-Neutral conditions was superior to that observed in the F-Neutral condition, overall. No significant interactions emerged from this analysis.

Error Analysis

To further explore the nature of the LHD subjects' impairment in perceiving emphatic stress cues, the data from the three Focus Perception tasks were examined collectively for the presence of a differential response pattern reflected in each group's errors. Figure 25 represents the percentage of stimuli designated by each group as exhibiting the correct location of focus, "no focus", or incorrect focus location (i.e., a "placement" error), by emotion type.

As may be seen, the majority of errors committed by the NC subjects involved misperceiving focused items as exhibiting "no focus", and very seldom involved placement errors, whereby the location of emphasis within the utterance was interchanged (i.e., substituting sentence-initial for sentence-final focus, or vice versa). A similar error pattern was evident for the RHD group, although the RHD patients tended to make a higher percentage of placement errors than the NC group. By contrast, the LHD group was far more likely to commit placement errors than the other two groups, frequently misidentifying the position of emphasis cues within the utterance. Nonetheless, a response of "no focus" remained the most frequent type of error response observed for the LHD subjects overall. In general, it was noted that focus expressed in sad utterances was

Figure 25. Percentage of total stimuli identified by each group as correct, as a "placement" error, or as lacking emphasis ("neutral").



particularly susceptible to being labelled as exhibiting "no focus" by all three groups, consistent with the observation that accuracy tended to be lowest for this particular sentence type.

2) Emotion Perception

As outlined above, recognition of the vocal cues to *emotional* distinctions was tested in the same subject sample, employing a highly similar design, including the same set of base stimuli presented in the focus perception tasks. The accuracy of the three diagnostic groups in labelling the emotional meaning of verbal stimuli was examined using three 3 X 3 X 2 X 3 mixed factorial design ANOVAs, performed separately for each emotion perception task (Baseline, D-Neutral, F-Neutral). Group membership (NC, RHD, LHD) was the between-groups factor for each analysis, and EMOTION (angry, sad, happy), MODALITY (declarative, interrogative), and FOCUS (initial, final, none) served as within-subjects factors. As before, post hoc inspection of significant findings was accomplished using Tukey's method ($p<.05$), wherever appropriate.

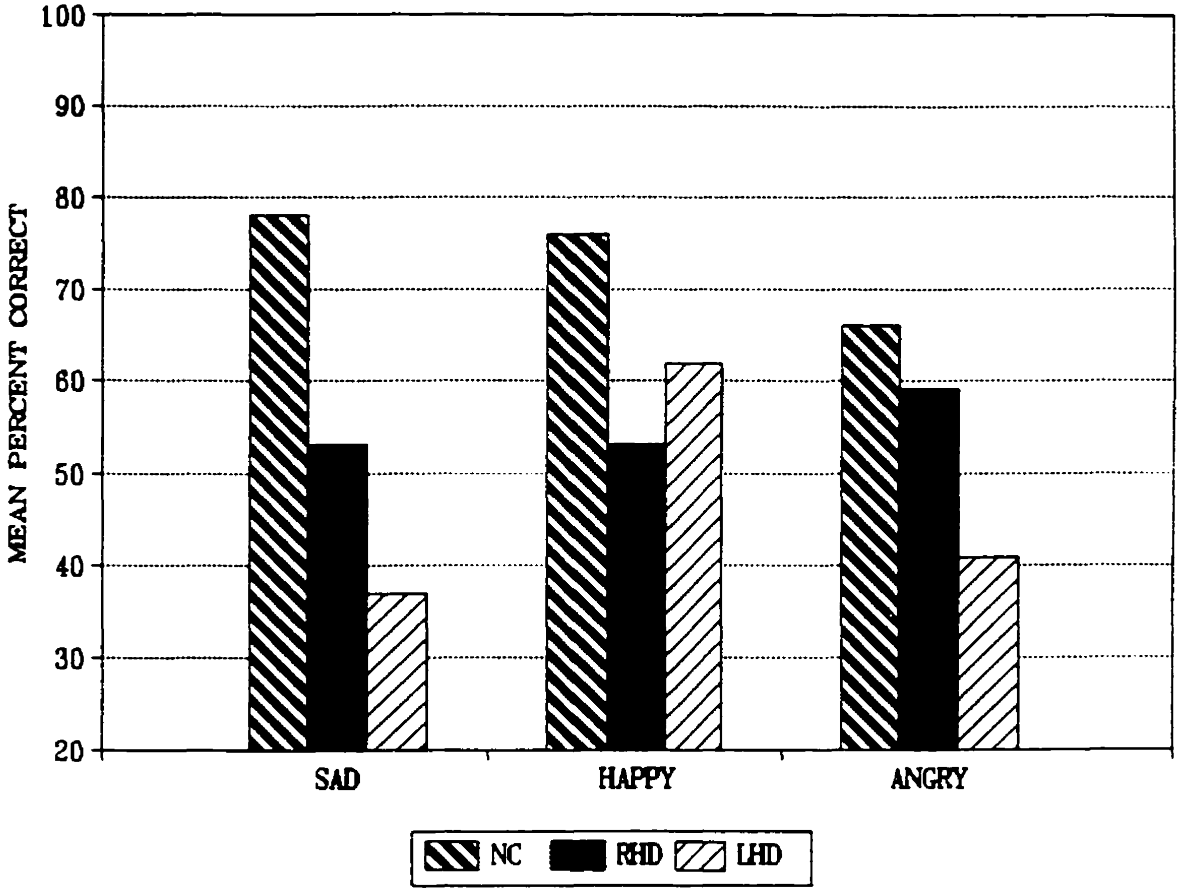
(i) <u>Baseline</u>

The Baseline task sought to establish each group's proficiency at deriving the emotional significance of vocal cues in simple utterances, and explored whether the identification of emotional prosody varies as a function of concurrent linguistic prosodic content (i.e., differences in focus distribution, sentence modality). Mean accuracy scores in identifying sad, happy, and angry stimuli on the Baseline task are presented for each group in Table 22, as a function of focus location and linguistic modality.

The ANOVA (GROUP X EMOTION X MODALITY X FOCUS) performed on these data uncovered a significant main effect for GROUP [$F(2,27) = 7.51, p<.01$]. Post hoc exploration of this effect indicated that the NC subjects were significantly better at identifying the emotional meaning of prosodic patterns, overall, than both the RHD and LHD subjects (who did not significantly differ). A significant GROUP X EMOTION interaction [$F(4,54) = 4.43, p<.01$], displayed in Figure 26, revealed that the performance of the LHD subjects on the Baseline task varied as a function of emotion type, their accuracy in labelling happy prosody surpassing that for both angry and sad prosody. The performance of the RHD and NC subjects did not significantly differ as a function of emotion type on this task.

Significant main effects for EMOTION [$F(2,54) = 3.23, p<.05$] and MODALITY [$F(1,27) = 29.35, p<.001$], and significant interactions of EMOTION X FOCUS [$F(4,108) = 4.62, p<.01$] and EMOTION X FOCUS X MODALITY [$F(4,108) = 2.80, p<.05$] were also noted. Inspection of the three-way interaction considered the influence of sentence modality and focus distinctions on the perception of each emotion. For sad and happy prosody, no differences in perceptual accuracy were observed as a function of sentential focus or linguistic modality. By contrast, the identification of angry prosody was shown

Figure 26. Interaction of group and emotion on the perception of emotional prosody on the Baseline task.



to be influenced by concurrent alterations in prosodic content; anger was interpreted significantly better in declarative utterances than interrogative utterances when no focus was present, and accuracy for angry prosody was higher in declarative utterances without focus relative to matched declarative utterances with sentence-initial focus.

Table 22. Percent correct responses (by group) in identifying sad, happy, and angry stimuli on the Baseline task, as a function of focus location and sentence modality.

		PERCENT CORRECT								
		NC			RHD			LHD		
Modal.	*Focus*	*S*	*H*	*A*	*S*	*H*	*A*	*S*	*H*	*A*
Declar.	None	78	66	92	51	64	73	38	65	53
	Initial	86	82	60	67	53	67	36	60	36
	Final	84	88	80	49	60	62	38	71	47
Inter.	None	82	72	58	47	51	49	42	64	35
	Initial	76	74	40	58	44	42	35	56	35
	Final	64	72	68	49	47	60	35	56	40
MEAN (SD)		78 (8)	76 (8)	66 (18)	54 (8)	53 (8)	59 (11)	37 (3)	62 (6)	41 (7)

Note: S = Sad, H = Happy, A = Angry, (*) indicates performance expected by chance for a four alternative forced choice paradigm based on 95% confidence limits, where chance=25% ± 9 (binomial distribution).

(ii) Neutral Duration (D-Neutral)

To test the importance of temporal acoustic markers on the perception of emotional prosody by brain-damaged and non-brain-damaged adults, the duration of vocalic segments of emotional stimuli were "neutralized" (i.e., rendered equivalent to those observed in matched neutral stimuli) prior to presentation in the D-Neutral task. Table 23 summarizes the outcome of this task, presenting the mean identification scores for sad, happy, and angry stimuli for each of the three diagnostic groups, as a function of sentence modality and emphasis location within the utterance.

It is clear from Table 23 that the manipulation of temporal parameters of emotional stimuli on the D-Neutral task exerted limited (although some) influence on the ability of the three groups to identify vocal affective meanings relative to their performance on the Baseline task. The ANOVA performed on these data revealed a significant main effect for GROUP [$F(2,27) = 8.25, p<.01$]. Follow-up tests established that, when emotional stimuli lacked appropriate temporal cues, both the RHD and LHD groups (which did not significantly differ) were less able to interpret the meaning of emotional prosody than the NC group. A significant main effect for MODALITY [$F(1,27) = 42.37, p<.001$], and significant interactions of EMOTION X FOCUS [$F(4,108) = 3.46, p=.01$], EMOTION X MODALITY X FOCUS [$F(4,108) = 3.38, p=.01$], and a four-way interaction of GROUP X EMOTION X MODALITY X FOCUS [$F(8,108) = 2.40, p<.05$] were also produced. The effects of sentence modality and focus location on each group's performance is presented for sad, happy, and angry prosody in Figures 27 (a-c), respectively.

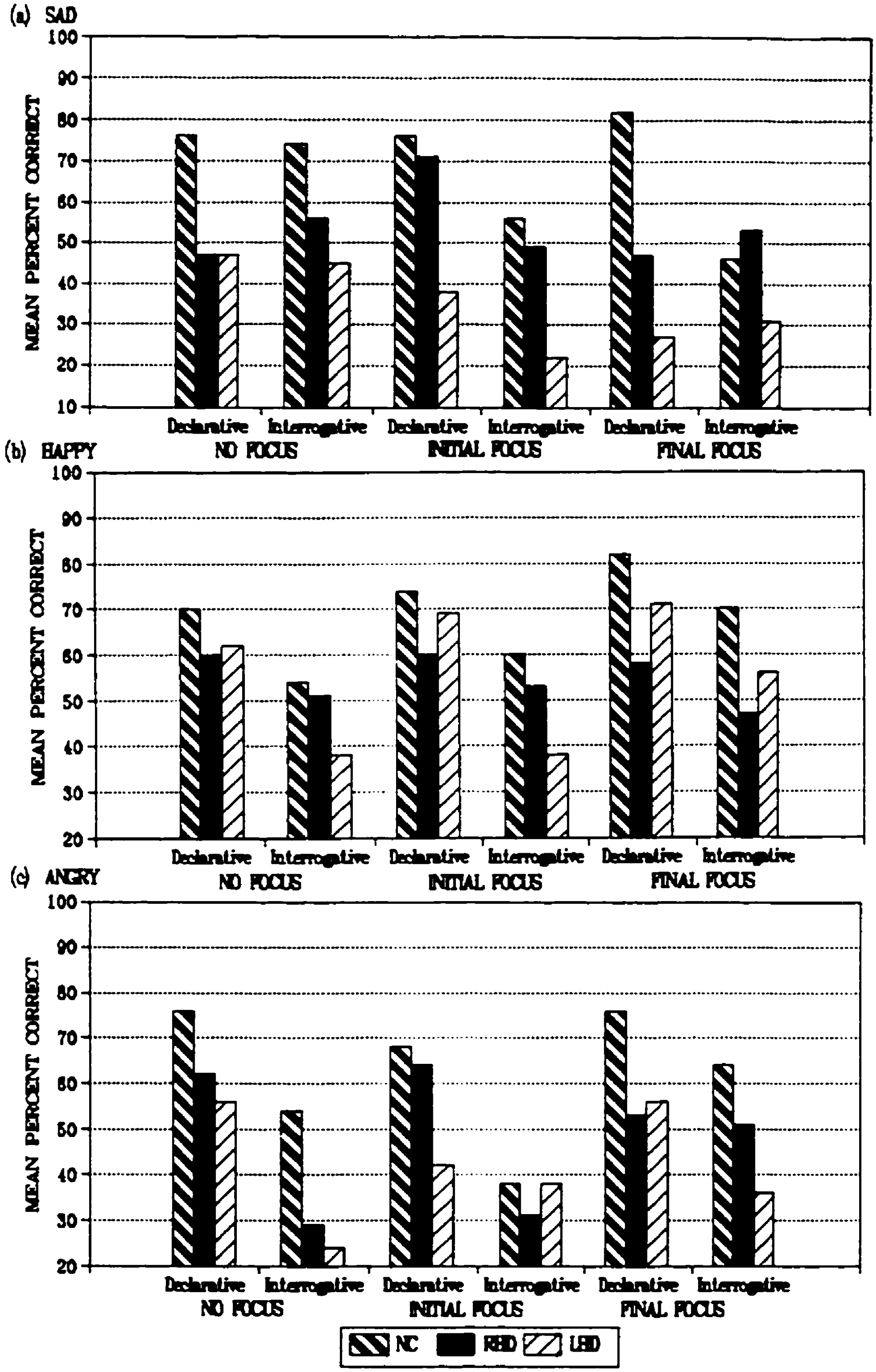
Table 23. Percent correct responses (by group) in identifying sad, happy, and angry stimuli on the D-Neutral task, as a function of focus location and sentence modality.

		PERCENT CORRECT								
		NC			RHD			LHD		
Modal.	*Focus*	*S*	*H*	*A*	*S*	*H*	*A*	*S*	*H*	*A*
Declar.	None	76	70	76	47	60	62	47	62	56
	Initial	76	74	68	71	60	64	38	69	42
	Final	82	82	76	47	58	53	27*	71	56
Inter.	None	74	54	54	56	51	29*	45	38	24*
	Initial	56	60	38	49	53	31*	22*	38	38
	Final	46	70	64	53	47	51	31*	56	36
MEAN (SD)		68 (14)	68 (10)	63 (15)	54 (9)	55 (5)	48 (15)	35 (10)	56 (15)	42 (12)

Note: S = Sad, H = Happy, A = Angry; (*) indicates performance expected by chance for a four alternative forced choice paradigm based on 95% confidence limits, where chance=25% ± 9 (binomial distribution).

Pairwise comparisons among the means of the four-way interaction examined the influence of emotion type, sentence modality, and focus location on the accuracy of each individual subject group. For both the NC and RHD groups, it was shown that accuracy in labelling the three emotions did not differ as a function of the other two factors (focus, modality). However, for the LHD group, recognition of happy intonation was shown to be superior to that of the other two emotions under certain conditions; namely, identification of happy prosody surpassed that of sad prosody (sentence-initial and sentence-final focus) and angry

Figure 27. Interaction of group, emotion, sentence modality, and focus on the perception of (a) sad, (b) happy, and (c) angry prosody on the D-Neutral task.



prosody (sentence-initial focus only) when declaratives were presented. No differences among the three emotions were noted for the LHD group when interrogative sentences were presented or when utterances lacked sentential focus. This pattern suggests that the presence of emphatic stress cues in declarative stimuli somehow facilitated recognition of happy prosody for the LHD subjects (review Figure 27b).

Finally, for all three groups, the ability to identify angry intonation was again shown to be systematically influenced by differences in focus location and linguistic modality: both the NC and RHD subjects recognized anger less often in interrogatives than declaratives when sentence-initial emphasis was present, and both the RHD and LHD patients recognized anger less frequently in interrogatives than declaratives when no focus was present. A similar pattern was reported in the Baseline task, suggesting that interrogative intonation was less conducive to effective recognition of angry prosody, except when emphasis was located sentence-terminally (review Figure 27c). Other significant comparisons indicated that, for the NC group, sad intonation was recognized less reliably in interrogatives with sentence-final emphasis than matched declaratives, and for the LHD group, happy prosody was misperceived more often in interrogatives with initial focus than matched declaratives.

(iii) Neutral F_0 (F-Neutral)

The modulation of fundamental frequency in speech is thought to be critically associated with emotional expression and the perception of emotional attributes. Accordingly, the F-Neutral task sought to determine the extent to which unilaterally brain-damaged patients and healthy control subjects could identify discrete emotional messages in utterances devoid of these natural pitch distinctions (i.e., in utterances in which pitch contours reflected those found in matched "neutral" utterances). Mean accuracy scores for the identification of sad, happy, and angry prosody by each of the three subject groups are summarized in Table 24, by modality and focus type.

As may be seen, accuracy on the F-Neutral task was generally low, and the overall accuracy of the three subject groups in identifying emotional prosody differed by a notably smaller margin on this task (NC=45%, RHD=39%, LHD=30%). Statistical inspection of these data produced a significant main effect for GROUP [$F(2,27) = 4.45, p < .05$]. Follow-up analyses indicated that the NC group identified emotional prosody significantly better than the LHD group on the F-Neutral task overall. No significant differences were noted between the overall accuracy of the RHD group and either the NC or LHD groups; as was also suggested by the focus perception data, the emergence of fewer group differences on the F-Neutral task may have reflected the greater extent to which the NC subjects, and not the brain-damaged patients, were affected by the manipulation of F_0. No interactions with the group factor were produced by this analysis, contrary to the findings of the other two emotion perception tasks, where the LHD patients alone demonstrated a bias for happy prosody over sad and angry prosody.

Table 24. Percent correct responses (by group) in identifying sad, happy, and angry stimuli on the F-Neutral task, as a function of focus location and sentence modality.

		PERCENT CORRECT								
		NC			RHD			LHD		
Modal.	*Focus*	*S*	*H*	*A*	*S*	*H*	*A*	*S*	*H*	*A*
Declar.	None	60	12*	54	47	20*	49	25*	27*	44
	Initial	58	18*	40	36	29*	42	27*	20*	29*
	Final	74	24*	72	49	29*	51	31*	29*	35
Inter.	None	60	60	28*	42	51	27*	25*	42	18*
	Initial	38	34*	22*	33*	44	29*	13*	38	22*
	Final	32*	54	66	27*	44	56	27*	51	40
MEAN (SD)		54 (16)	34 (20)	47 (20)	39 (9)	36 (12)	42 (12)	25* (6)	35 (11)	31* (10)

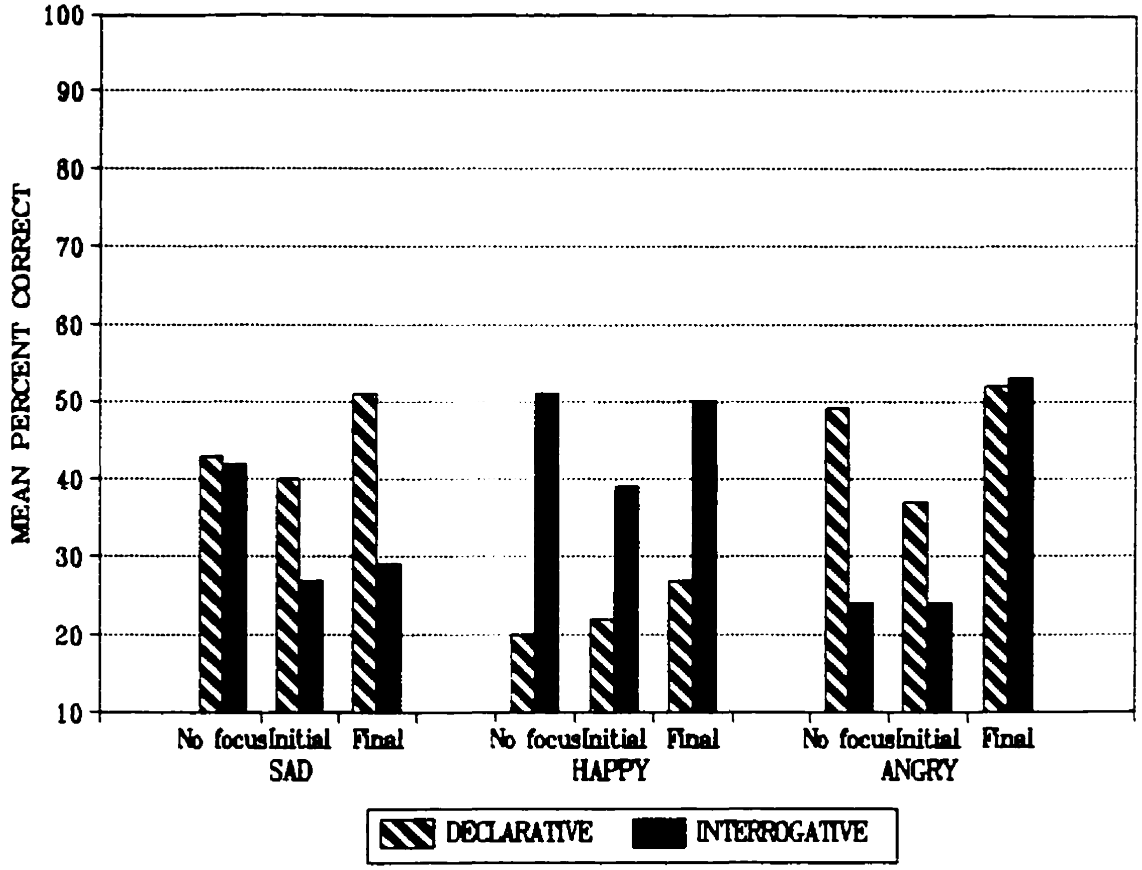
Note: S = Sad, H = Happy, A = Angry; (*) indicates performance expected by chance for a four alternative forced choice paradigm based on 95% confidence limits, where chance=25% ± 9 (binomial distribution).

A significant FOCUS main effect [$F(2,54) = 12.67, p<.001$], and significant interactive effects of EMOTION X MODALITY [$F(2,54) = 12.76, p<.001$], EMOTION X FOCUS [$F(4,108) = 4.21, p<.01$], and EMOTION X MODALITY X FOCUS [$F(4,108) = 7.15, p<.001$] emerged. Examination of the three-way interaction, illustrated in Figure 28, considered how differences in sentence modality and focus location influenced the recognition of vocal cues to each individual emotion in the absence of appropriate pitch modulation. For sad stimuli, no differences in accuracy were observed as a function of modality or focus location. For happy stimuli, recognition was significantly improved when presented in conjunction with interrogative intonation rather than declarative intonation for all focus types; this interaction was uniquely observed in the F-Neutral task. Finally, for angry stimuli, a pattern resembling that observed in the Baseline and D-Neutral tasks emerged: recognition of angry prosody was significantly lower for interrogative utterances than declaratives utterances when focus was absent; recognition of angry prosody was significantly higher for interrogative utterances when sentence-*final* emphasis was present than when sentence-initial or no emphasis was present.

Summary

In summary, the identification of emotional prosody was impaired in both LHD (3/3 tasks) and RHD (2/3 tasks) patients relative to non-neurological control subjects as assessed by three distinct emotion perception tasks. On the F-Neutral task, the accuracy of the NC group fell to that of the RHD group, leading to non-significant differences between the NC and RHD groups. Emotional recognition scores did not significantly differ between the RHD and LHD groups for any of the three emotion perception tasks, overall. An emotional bias was observed

Figure 28. Interaction of emotion, sentence modality, and focus on the perception of emotional prosody on the F-Neutral task.



uniquely for the LHD group, who displayed superior recognition of happy prosody on two of the three tasks (Baseline, D-Neutral).

The influence of sentence modality and sentence focus distinctions on the perception of discrete affects was most prevalent for angry prosody. The perceptual markers of interrogation, in the *absence* of sentence-final emphasis, generally interfered with the recognition of angry prosody by all three subject groups, irrespective of the acoustic composition of the stimuli. Systematic effects of sentence modality and focus distribution on the other two emotions were far less evident, although it was noted that accuracy for happy prosody was greater in interrogative than declarative utterances subsequent to the manipulation of F_0 parameters of the stimuli.

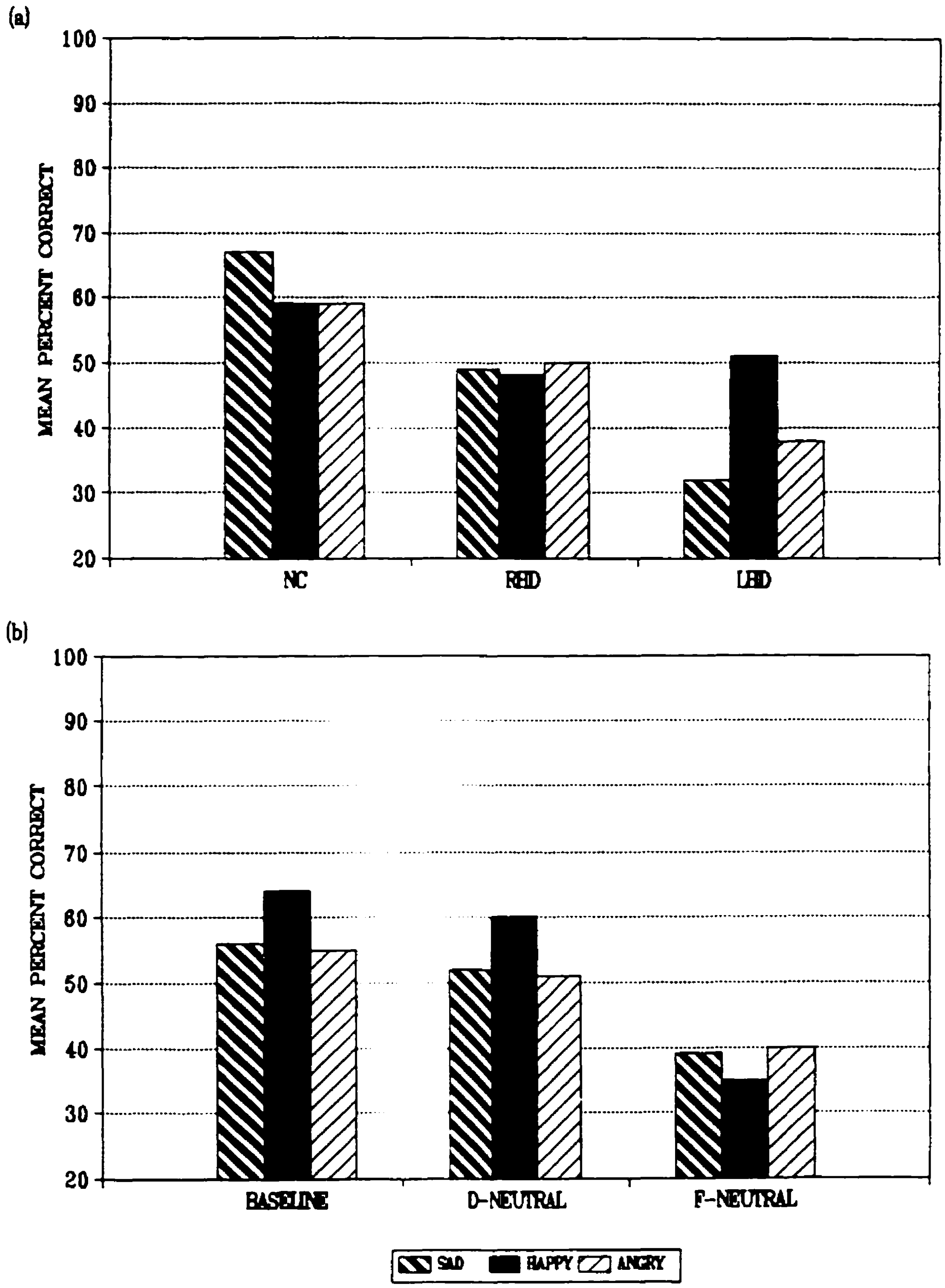
Effect of the Acoustic Manipulation of Emotional Cues: TASK Factor

The direct influence of manipulating temporal or spectral aspects of speech on the identification of emotional prosody by the three subject groups was examined using a 3 X 3 X 3 mixed ANOVA in a full factorial design. Group membership (NC, RHD, LHD) served as the between-subjects factor in this analysis, and EMOTION (sad, happy, angry) and TASK (Baseline, D-Neutral, F-Neutral) served as the within-subjects factors. The data collected for each task were collapsed across the two levels of "modality" and the three levels of "focus" prior to statistical inspection.

A significant main effect for GROUP was observed [$F(2,27) = 8.27$, $p<.01$]; this effect reflected a significant decrement in the ability of the RHD and LHD patients to identify emotional prosody relative to NC subjects overall. A significant GROUP X EMOTION interaction, represented in Figure 29a, also emerged [$F(4,54) = 3.31$, $p<.05$]. Pairwise comparisons conducted for each group individually indicated that only the accuracy of the LHD group differed as a function of the three emotions, with LHD subjects demonstrating significantly better recognition of happy stimuli than sad stimuli; (accuracy for angry stimuli did not differ from that of happy or sad stimuli for the LHD group). Comparisons of group accuracy for each individual emotion type indicated that the NC subjects were significantly better than the LHD subjects in identifying both sad and angry stimuli, and the NC subjects were significantly better than the RHD subjects in labelling sad stimuli. No significant between-group differences in the identification of happy prosody were revealed (review Figure 29a).

The emergence of a significant main effect for TASK [$F(2,54) = 60.74$, $p<.001$] was explained by the lower accuracy of all subjects in the F-Neutral task relative to both the Baseline and D-Neutral tasks. A significant EMOTION X TASK interaction, depicted graphically in Figure 29b, was further produced [$F(4,108) = 4.22$, $p<.01$]. Examination of this interaction indicated that all three emotions were identified more poorly in the absence of appropriate pitch modulation (F-Neutral task) relative to both the Baseline and D-Neutral tasks.

Figure 29. Interaction of (a) group and emotion, and (b) emotion and task in the Emotion Perception condition overall.

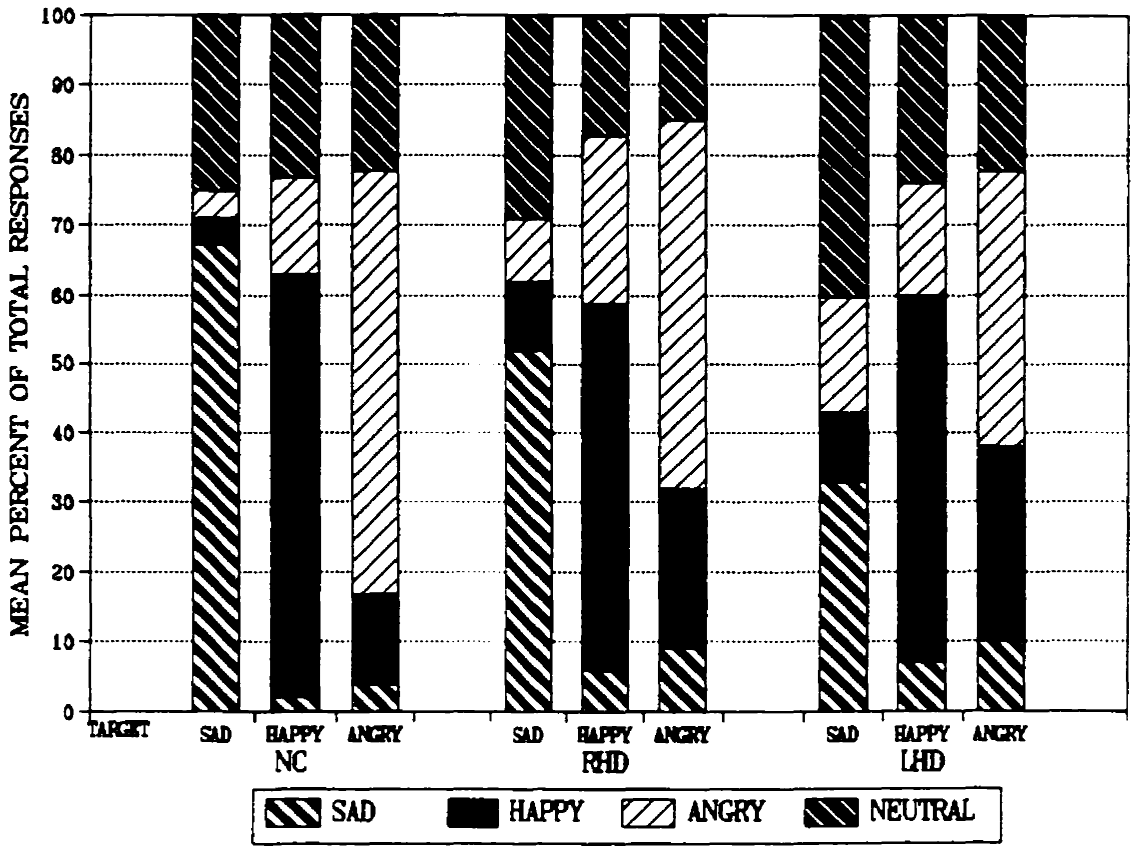


Accuracy for the three emotions did not differ significantly for any of the three tasks, overall.

Error Analysis

Errors in the identification of sad, happy, and angry prosody were subsequently examined to determine whether the response patterns of the three groups could be distinguished qualitatively; Figure 30 provides these data, collapsed across sentence modality, focus position, and task level. As may be seen, when errors occurred, happy and angry prosody were either interchanged, or designated as "neutral" by all subjects, very seldom being confused with "sad". Sad prosody was typically misidentified as "neutral" by all subjects when errors were observed; indeed, for the LHD group, a "neutral" response occurred more frequently than correct identifications of sad prosody, overall. Finally, the RHD and LHD patients substituted sad prosody for happy or angry prosody far more often than the NC subjects, who almost never committed this type of error. Despite these differences, the response patterns of the three groups were relatively comparable, and did not suggest marked qualitative differences.

Figure 30. Percentage of stimuli identified by each group as conveying sad, happy, angry, or "neutral" prosody.



DISCUSSION

Much remains unclear about how prosodic aspects of speech are translated into affective and non-affective messages in normal speech perception. Even less understood are the effects of unilateral brain injury, and possible compensatory mechanisms, on these auditory-perceptual processes in particular, and communicative competence in general. More exhaustively than in previous studies, Experiment 2 tested the capacity of right- and left-hemisphere-damaged patients to perceive prosodic forms of both an affective and non-affective nature. The presentation of auditory stimuli enriched with multiple prosodic cues, and the manipulation of specific temporal or spectral properties of the stimuli, were viewed as an effectual means of assessing the impact of unilateral brain injury on prosodic perception under a variety of conditions. Data from this experiment are further important in illuminating some of the normal processes underlying the perception of linguistic and emotional prosody, and the possible interaction of these two "types" of representations.

As noted earlier, the perception of emphasis is often critical to recognizing the propositional intent of speech, highlighting the semantic value accorded to individual constituents within an utterance. In the present study, three tasks assessed the ability of RHD, LHD, and NC subjects to locate the position of emphatic stress within short sentences; on each task, the LHD aphasic patients were significantly impaired relative to matched RHD patients, who performed at a comparable level to non-neurological control subjects in each case. The LHD patients were also impaired relative to the normal subjects on two of the three focus perception tasks (Baseline, D-Neutral), and exhibited defective recognition of emphasis cues relative to both the RHD and NC subjects when the data were collapsed across the three task levels. These findings demonstrate that patients with left hemisphere dysfunction and aphasia are disturbed in the capacity to perceive emphatic stress distinctions in speech (Kimelman, 1991), providing compelling evidence of a *left* hemisphere substrate for the perceptual mechanisms underlying these processes. This outcome is in accord with previously published data (Baum et al., 1982) and receives support from recently-conducted research on emphatic stress perception by aphasic adults (Baum, under review).

The observation that focus perception was spared subsequent to unilateral right-hemiphere insult is at odds with some previously-reported findings (Brådvik et al., 1991; Bryan, 1989; Weintraub et al., 1981). However, methodological inconsistencies and the use of less rigorous testing paradigms in previous reports may explain the different patterns noted. For example, two of the studies demonstrating defective recognition of emphatic stress in RHD patients (Brådvik et al., 1991; Bryan, 1989) presented no more than 12 trials to their subjects for perceptual identification. In contrast, the overall pattern reported in the current experiment was based on the results of three distinct tasks of 80 trials each, a far more reliable sample. With respect to the investigation conducted by Weintraub and her colleagues (1981), those authors assessed the success of RHD patients in *discriminating* (i.e., making a same/different judgement about) paired utterances differing in emphatic stress location, and within a single analysis, assessed the ability of RHD patients to differentiate statement/question intonational contrasts as well. Thus, due to methodological considerations, the extent to which the RHD patients reported by Weintraub et al. (1981) were receptively impaired for emphatic stress cues or for linguistic intonation--a deficit reported elsewhere in the literature (Brådvik et al., 1991; Heilman et al., 1984; Pell and Baum, 1997a)--is unclear. These inconsistencies render the findings of previous investigations on emphatic stress less certain, favouring the present conclusion that RHD patients are largely unimpaired in the receptive processing of emphatic stress cues in speech (Baum, under review).

The stimuli presented for focus recognition were also presented to the same subject sample to test their recognition of *emotional*-prosodic attributes. Results obtained for the three tasks of emotion perception indicated a somewhat different pattern of group performance from that reported in the focus perception condition. Namely, both the RHD and LHD patients performed at an inferior level when compared to the control subjects on all emotion perception tasks except for the F-Neutral task, where the accuracy of the RHD and NC subjects did not significantly differ; (exceptional patterns observed in the F-Neutral data are discussed below). Both clinical groups also exhibited disturbed *overall* recognition of emotional prosody relative to the normal subjects when the data were collapsed across the three task levels. A disruption of emotional prosody perception following both left- and right-hemisphere insult is in accordance with much recently-published data (Cancelliere & Kertesz, 1990; Darby, 1993; Starkstein et al., 1994; Van Lancker & Sidtis, 1992), arguing that *both* the left and right hemispheres of the brain may normally be engaged in decoding emotional attributes of speech.

The comparable performance of the RHD and LHD patients in the emotion perception condition stands in sharp contrast to the performance of these two groups in the focus perception condition, where the accuracy of the RHD patients *always* exceeded that of the LHD patients. Comparison of group performance across the focus and emotion conditions, therefore, suggests that right hemisphere lesions may be specifically tied to a receptive disruption for emotional, but not linguistic, aspects of prosody in speech perception (Emmorey, 1987; Heilman et al., 1984; Ross, 1981). This finding is compatible with the notion that right hemisphere mechanisms are essential in decoding communicative stimuli of an emotional nature (e.g., Borod, 1992), and more generally, that hemispheric specialization for prosody may somehow depend on the functional (i.e., linguistic vs. affective) valence of prosodic features in speech perception (Behrens, 1985; Emmorey, 1987; Van Lancker, 1980).

However, as the RHD and LHD samples tested in the current study displayed comparable deficits in the recognition of emotional prosody, the data reported herein are discordant with accounts claiming *exclusive* right-hemisphere control of emotional prosody (Blonder et al., 1991; Bowers et al., 1987; Ehlers &

Dalby, 1987; Heilman et al., 1975; Hughes et al., 1983; Ley & Bryden, 1982; Ross, 1981; Tucker et al., 1977). As discussed in the introduction, the absence of a LHD subject group in many previous investigations of the identification of emotional-prosodic meanings (Heilman et al., 1975; Hughes et al., 1983; Ross, 1981; Tucker et al., 1977) may have led to premature claims about the right hemisphere's exclusive role in decoding vocal affect. The present data mitigate against those authors' view that only right-hemisphere lesions yield deficits in the recognition of emotional prosody, demonstrating that lesions confined to either cerebral hemisphere may underlie such impairments (Cancelliere & Kertesz, 1990; Darby, 1993; Starkstein et al., 1994; Van Lancker & Sidtis, 1992). This is not to say that the precise manner in which right and left hemisphere insult disturbs emotional prosody recognition does not differ, an issue to be addressed in future investigations.

The fact that fewer group differences in overall accuracy emerged on the F-Neutral task (where normally-occurring pitch cues were rendered unavailable) within both the focus and emotion perception conditions merits some comment. As summarized above, on the F-Neutral task, the accuracy of the normal subjects did not significantly differ from that of the LHD and RHD patients in the focus and emotion conditions, respectively. Although a significant group by task interaction did not emerge in the statistical analyses (i.e., the three groups were not differentially influenced by the acoustic manipulation of the stimuli in either the focus or emotion condition), trends in the data suggest that the NC subjects may have been somewhat more affected by the absence of normal pitch modulation on the F-Neutral task than either the LHD or RHD subjects. The NC subjects' accuracy on the F-Neutral task (across conditions) illustrated a greater decrement (21%) than that of the RHD (16%) and LHD (10%) patients when compared to accuracy on the Baseline task: this trend may have caused the accuracy of the NC group to approximate that of the patient groups on the two F-Neutral tasks, resulting in fewer group distinctions. Thus, although pitch fluctuation was clearly of primary importance to all three groups, normal subjects may have been particularly reliant on pitch change in perceiving prosodic distinctions related to emphasis and emotional content (see also Baum, under review).

The emergence of distinct group patterns as a function of individual stimulus types in both perception conditions points to further differences in the underlying perceptual mechanisms employed by brain-damaged and non-brain-damaged subjects. These differences were especially noteworthy in the focus perception condition. For example, requiring subjects to rely solely on pitch fluctuation when perceiving focus (D-Neutral task) selectively disturbed the two patient groups relative to the normal subjects, but only for those stimuli in which pitch fluctuation was minimal (i.e., sad utterances). This finding suggests that both the RHD and LHD patients were less efficient than normal listeners at processing the pitch attributes of emphatic stress when pitch distinctions in the signal were least informative of the target response. Moreover, it implies that both clinical groups derived benefit from the temporal correlates of emphasis when presented with sad stimuli on the Baseline and F-Neutral tasks, where this selective pattern of disruption was not observed.

Relatedly, requiring subjects to rely solely on temporal parameters when perceiving emphasis (F-Neutral task) resulted in a selective deficit for the LHD patients relative to the RHD patients, but only when *declarative* contours were presented. Inspection of the acoustic data suggests that, despite neutralizing the magnitude of pitch change associated with focused items on the F-Neutral task, interrogative stimuli remained marked by the *direction* of pitch change occurring on focused items (i.e., pitch accents displayed a rising contour on focused items within interrogatives sentences but a falling contour within declarative sentences, Eady & Cooper, 1986; see Experiment 1). These categorical alterations in the direction of the pitch accent contained in interrogative contours may have proven sufficient for the LHD patients in determining emphasis location, their deficits on this task being confined to those stimuli in which pitch markers were completely absent and only temporal cues to emphasis could be harnassed (i.e., declarative contours). This result indicates that LHD patients may demonstrate a subtle deficit in processing temporal parameters of linguistic-prosodic stimuli, a shortcoming largely obscured when pitch markers are salient (Van Lancker & Sidtis, 1992). Collectively, the sensitivity of the brain-damaged patients to specific stimulus types in the face of the acoustic manipulations affirms that RHD and LHD patients normally require both spectral *and* temporal aspects of prosody when processing emphatic stress distinctions, redundancy not always required by the normal system on such a task.

Finally, in the emotion perception condition, the LHD group showed a bias towards recognition of happy prosody over sad and angry prosody on all tasks except the F-Neutral task. Such a bias has been noted previously using a comparable stimulus set (Pell & Baum, 1997a) and may underscore the proficiency of the aphasic patients at processing prosodic forms displaying large excursions in continuous F_0 features (Pell & Baum, 1997b). Alternatively, in view of the emotional responses permitted in the present experiment, recognition of happy prosody may have been enhanced by a predilection for "positive" stimuli in the LHD patients (Ahern & Schwartz, 1979; Sackeim, Putz, Vingiano, Coleman & McElhiney, 1988). However, this latter explanation is unlikely given the weight of evidence opposing the "valence" hypothesis (Ehlers & Dalby, 1987; Heilman et al., 1984; Pell & Baum, 1997a; 1997b), including the failure of the present data to produce the converse pattern (i.e., RHD patients demonstrating a perceptual bias for "negative" affects), also predicted by this hypothesis. The probability that the pitch attributes of happy stimuli guided the LHD patients' performance is reinforced by the lack of an emotional bias on the F-Neutral task, where F_0 properties of the stimuli were rendered "neutral" and therefore ineffective determinants of emotional content.

More generally, results obtained herein demonstrate that for both brain-damaged and non-brain-damaged subjects, the perception of linguistic prominence and emotional prosody varies as a function of concurrent specifications in intonation contours. This interaction was particularly evident in the focus perception condition, where all subjects displayed superior recognition of emphatic stress in utterances spoken in a happy or angry tone than those spoken in a sad tone. Emphasis detection in "neutral" utterances was generally poorer than that in happy or angry utterances, although frequently better than in sad utterances. These data point to a perceptual interaction between the cues that signal linguistic prominence and those underlying different affective modes, a finding heretofore unreported. Perhaps more significantly, these results underscore the sensitivity of both normal subjects and unilateral left- and right-brain-damaged patients to the perceptual attributes of specific emotional patterns while performing a *linguistic* task, as measured in the focus perception condition.

As noted earlier, the focus perception stimuli differed markedly as a function of emotion when acoustic parameters of the stimuli were examined. Perhaps most importantly, the magnitude of F_0 change associated with focused items in both sentence-initial and sentence-final position was extremely large for happy and angry stimuli and extremely small for sad stimuli when compared to neutral stimuli. These highly exaggerated pitch changes observed on focused items in happy and angry stimuli may have enhanced the prosodic contrast between emphasized and unemphasized segments within these contours, resulting in improved perceptual acuity and higher accuracy scores for happy and angry stimuli when compared to stimuli displaying relatively small pitch accents (i.e., sad stimuli). It is unlikely that temporal differences across emotional categories facilitated emphasis perception, as focused vowels in sentence-initial and sentence-final positions were actually longest (i.e., presumably most indicative of focus) for the sad stimuli, for which recognition of focus was *least* accurate.

Evidence of a facilitative relationship between "exaggerated" prosodic patterns and accuracy on speech perception tasks is of potential importance to aphasic patients with auditory comprehension deficits. Based on previous research, it has been shown that narrative comprehension may be significantly enhanced in aphasic listeners when "exaggerated stress" is placed on key target words within the narrative (Pashek & Brookshire, 1982; Kimelman & McNeil, 1987), and that prosodic modulation *preceding* stressed target words may further improve word comprehension by aphasic adults on certain tasks (Kimelman & McNeil, 1989). Thus, despite present indications that LHD patients are impaired when required to perceptually *isolate* emphatic cues within the speech stream (see also Baum et al., 1982; Kimelman, 1991), other data suggest that aphasic patients continue to derive some benefit from emphatic specifications occurring throughout the utterance (or on key words within a narrative) when auditory comprehension of speech is tested.

Implications in the present data that affective modes which naturally enhance or "exaggerate" distinctions in linguistic-prosodic content render emphatic items more salient to all listeners, including both fluent and nonfluent aphasics, serve to extend these previous findings (Pashek & Brookshire, 1982; Kimelman &

McNeil, 1987; Kimelman & McNeil, 1989). As suggested by Fernald's (1984; 1989) data on motherese, the more "informative" nature of expanded contours probably renders the communicative intent of the speaker more intelligible to the listener, leading to the use of such forms in speech directed to infants or foreign language speakers. Although a link between exaggerated or expanded intonation contours and improved sentence comprehension abilities in aphasic adults remains highly speculative at present, this line of inquiry nonetheless presents a promising direction for future research.

In addition to emotional distinctions, differences in linguistic modality also influenced the ability to perceive focus location, all subjects displaying superior recognition of emphasis in sentence-final position when interrogative intonation was presented. Examination of the acoustic data indicated that a substantial rise in terminal F_0 served as the principle marker of interrogation as well as the primary physical determinant of linguistic focus. For the present stimulus set, therefore, the need to convey focus and interrogation coincided in sentence-final position, resulting in utterances with relatively little pitch modulation preceding the sentence-final item but an extremely large rise in pitch on the sentence-final item. The coincidence of these cues in sentence-final position probably rendered focused items especially salient to subjects for interrogative stimuli.

Finally, the perceptual identification of emotional prosody was relatively little influenced by the presence of concurrent linguistic-prosodic cues. Exceptionally, angry prosody was recognized poorly by subjects when presented as an interrogative without focus or with sentence-initial focus, across task levels. Examination of the acoustic data for the emotional stimuli did not readily indicate why recognition of angry prosody was uniquely dependent on changes in focus distribution and sentence modality. However, despite a rating procedure whereby the perceptual validity of the stimuli was established, comments from experimental subjects following the testing session indicated that a subset of the angry stimuli may have sounded closer to "surprise" intonation than to "angry", leading to some confusion for this emotional category. Possibly, the perception of a rise in pitch occurring relatively early in angry utterances--a pattern exclusively displayed by interrogative contours with initial or no focus--is more conducive to "surprise" intonation, all else remaining equal. Further inquiry into the perceptual correlates of discrete emotions may help resolve these issues.

In summary, the perception of prosodic distinctions by unilateral right- and left-brain-damaged subjects suggests that cerebral activation for the linguistic and emotional components of prosodic patterns may differ on several counts. Right hemisphere dysfunction was shown to selectively disturb *emotional* attributes of vocal cues in speech perception (these patients were not receptively impaired for emphatic stress). However, a disruption of emotional prosody was *not* unique to the RHD patients, as LHD aphasic patients also displayed difficulties in the emotion recognition condition, as well as the emphatic stress perception condition. The observation that certain emotional modes serve to enhance distinctions in linguistic-prosodic content for the listener, and more generally, that emotional and linguistic specifications interact in auditory perception, provide a basis for future research in this area.

CONCLUDING REMARKS

Many claims have been made about the right hemisphere's role in the production and perception of speech prosody. Despite the fragmentary nature of these claims, it is commonly held that right hemisphere mechanisms must somehow subserve these psycholinguistic functions, however ill-defined at present. The present study of both expressive and receptive abilities sought to understand how RHD patients were influenced by specific acoustic or perceptual correlates of prosodic patterns when engaged in verbal behaviour. Assessing the *interplay* of various prosodic features, although complex, was necessary to illuminate the RHD patients' performance in a context simulating that of spontaneous speech in many respects.

Certainly, the stimuli upon which the present conclusions are based were not fully representative of naturalistic speech samples, but rather, were constrained by the elicitation procedure utilized in both experiments. The possibility that "acting" ability influenced the acoustic measures obtained in the present study somewhat, especially in the production experiment (perceptual stimuli were produced by a single subject and then submitted to a rating procedure), cannot therefore be discounted. However, the close fit of the present data with previous *normal* acoustic descriptions of both read and spontaneous speech (e.g., Cooper et al., 1985; Eady & Cooper, 1986; Williams & Stevens, 1982) suggests that the elicitation procedure employed herein was indeed a suitable vehicle for investigating prosody in RHD adults. Moreover, the likelihood that inter-subject differences in the ability to simulate emotions (as well as other aspects of prosody) were equally prevalent in both the RHD and NC groups diminishes the potential contribution of this factor.

An attempt to correlate clinical attributes of the RHD patients described herein with their prosodic deficits, although potentially revealing, was beyond the purview of the present investigation. However, following comments made by Pell and Baum (1997a), it is noteworthy that expressive and receptive prosodic deficits uncovered in the current RHD sample coincided with other severe neurologic signs, such as hemispatial visual neglect in approximately half of the RHD sample. As hypothesized earlier, evidence of a lasting behavioural neglect in RHD patients with stable lesions may point to relatively extensive neurologic damage in these patients, a potential determinant of prosodic disruption (Pell & Baum, 1997a). In contrast, RHD patients with milder neurologic signs may present less frequently with expressive or receptive disturbances of speech prosody when compared to matched LHD or normal subjects (Pell & Baum, 1997a; Schlanger et al., 1976; Van Lancker & Sidtis, 1992). The complexity of the current experimental design did not permit an examination of individual patterns of performance within each subject group, preventing a test of this hypothesis within the present RHD sample. Nonetheless, a relationship between prosodic deficits and concurrent behavioural signs produced by unilateral right-hemisphere insult remains possible and awaits further analysis.

Despite examining hemispheric involvement for various prosodic functions, an in-depth appraisal of lesion localization within the right (or left) hemisphere was also beyond the goals of the current investigation. However, prior research has identified neuroanatomical regions of potential importance within the cerebral hemispheres that may be linked to disturbances of prosodic function. Specifically, a potential association between prosodic impairment and lesions localized to right and/or left temporoparietal regions of the cortex has been documented (Blonder et al., 1991; Heilman et al., 1975; Ross, 1981; Schlanger et al., 1976; Tucker et al., 1977; Van Lancker & Sidtis, 1992). More recently, *subcortical* damage involving the basal ganglia has been strongly implicated in producing prosodic disturbances in a number of neurogenic populations, including patients with focal right and left basal ganglia lesions (Blonder et al., 1989; Brådvik et al., 1991; Cancelliere & Kertesz, 1990; Cohen et al., 1994; Pell, 1996; Ross & Mesulam, 1979; Speedie et al., 1990; Starkstein et al., 1994; Van Lancker & Pachana, 1995). As lesion site varied somewhat within the clinical samples reported herein, the present data can neither substantiate nor disconfirm these hypotheses. Future investigations which aim to correlate circumscribed neuroanatomical regions within the right or left hemisphere of the brain with select prosodic deficits (such as that conducted by Cancelliere and Kertesz, 1990) are clearly necessary.

Finally, as is also true for segmental aspects of speech, hypotheses arrived at to explain the *production* data on prosody do not always correspond well with those accounting for patterns in the *perception* data on prosody. The present study sought to address several current hypothetical descriptions of how prosody is lateralized in the brain (Behrens, 1988; Ross, 1981; Van Lancker & Sidtis, 1992). With respect to the "functional load" hypothesis (e.g., Behrens, 1985; Emmorey, 1987; Ross, 1981), results obtained for the perception experiment were generally consistent with the predicted pattern: RHD patients exhibited deficits for only emotional attributes of speech, and of the three diagnostic groups, LHD patients were exclusively impaired for prosodic cues of a linguistic nature.[10] Although an alternative explanation was adopted to account for patterns in the production data, it is noteworthy that results obtained for Experiment 1 also conformed in a general (albeit partial) manner to those anticipated by the functional role hypothesis; namely, RHD patients exhibited greatest disturbance in producing long-term parameters of prosody important in *emotional* signaling. (The absence of a LHD group in Experiment 1--due to the difficulty of the production task--did not permit a direct test of the left hemisphere's role in prosody production.)

As discussed in detail in Experiment 1, expressive abnormalities in the RHD patients' speech were attributed to disturbed regulation of continuous, as opposed to categorical, aspects of prosodic stimuli. Given the apparent quantitative nature of this deficit and its hypothesized occurrence at the *implementation* level, such an interpretation was preferred over a functional account of the production data, despite the fact that a functional description was

---

[10]Note that the LHD patients were also impaired in recognizing emotional prosody, a pattern that cannot be explained within the functional role hypothesis.

adopted to explain patterns in the perception data. Nonetheless, as acknowledged earlier, the fact that continuous and categorical aspects of prosody are highly, if not inextricably, linked to the emotional and linguistic function of these cues in speech (e.g., Ladd et al., 1985) suggests that a functional interpretation of both the production and perception data reported herein cannot be discounted (Behrens, 1985; Emmorey, 1987).[11] However, it is equally possible that further investigation in this area will uncover evidence of hemispheric asymmetries for the continuous and categorical features of prosodic stimuli in the *receptive* as well as expressive mode, as suggested by previous data (Pell & Baum, 1997b).

Other current hypotheses in the literature on prosody derived limited support from the present investigation of both production and perception skills. For example, there were few indications in the present data that the right and left cerebral hemispheres of the brain are specialized for F_0 and duration parameters of prosodic stimuli, respectively (Robin et al., 1990; Van Lancker & Sidtis, 1992). For production abilities, both duration and F_0 features of prosody were shown to be aberrant in the RHD patients' speech when these parameters were used continuously. Furthermore, in perceiving emphatic stress and emotional distinctions, neither the RHD nor LHD groups was selectively influenced by the absence of only duration or F_0 cues to these meanings. The failure of the current

---

[11]Given the RHD patients' difficulty in regulating the magnitude of F_0 change underlying emphasis, the notion that emphasis is composed of both a "linguistic" and "affective" component (the latter representing the speaker's pragmatic decision to signal the *extent* of emphatic highlighting desired) must be accepted if a functional interpretation of the current production data is to be adopted.

study to support the "cue hypothesis" is in keeping with much recently-published data (Dykstra et al., 1995; Gandour et al., 1995; Pell & Baum, 1997b).

The contention that right hemisphere mechanisms contribute to the processing of relatively large prosodic units such as intonation contours, but are minimally implicated when processing small prosodic constituents such as linguistic stress (Behrens, 1989; Emmorey, 1987; Gandour et al., 1992), also receives little support from the present study. As discussed in Experiment 1, although long-term properties of intonation contours were most frequently impaired in the RHD speakers' utterances (as predicted by the "domain hypothesis"), success in producing short-term parameters of emphatic stress tokens depended on the nature of the cue to be modulated and was not entirely normal in the RHD group. The emergence of the pattern predicted by the "domain hypothesis" for the RHD patients in Experiment 2 (i.e., RHD patients were impaired in recognizing emotional prosody but not emphatic stress) is tempered by the observation that this pattern is simultaneously predicted by the *functional role* of these cues in speech perception, as noted earlier. Moreover, the comparable level of impairment displayed by the RHD and LHD patients in the emotion identification condition is incompatible with the notion that the right hemisphere *exclusively* contributes to intonational decoding.

In fact, as noted briefly above, the LHD patients' disturbance in identifying emotional prosody in Experiment 2 is not easily accomodated by either the domain hypothesis or the functional role hypothesis. Although not a novel finding

(Cancelliere & Kertesz, 1990; Schlanger et al., 1976; Seron et al., 1982; Tompkins & Flowers, 1985; Van Lancker & Sidtis, 1992), evidence that focal left-hemisphere lesions sometimes produce affective-prosodic deficits of a comparable scale to those witnessed in RHD patients has not received much attention in the prosody literature (although see Ross, Thompson & Yenkosky, 1997, for a recent study that attempts to address this issue). Accordingly, it remains to be determined whether difficulties experienced by LHD aphasic patients on emotional prosody tasks are of a similar nature to those hypothesized for RHD patients, or whether, for example, LHD patients are more susceptible to increased cognitive demands on such tasks (Tompkins & Flowers, 1985). Moreover, it is recognized that the need to elicit a verbal response or decision on most emotional tasks may selectively affect the performance of LHD (but not RHD) patients, but the ramifications of this verbal component on emotional recognition by aphasic patients are poorly understood. These uncertainties suggest that close attention to the factors underlying prosodic performance in both RHD *and* LHD patients may benefit future research in this area.

Remaining discrepancies in the production and perception data on prosody suggest the probability that prosodic abnormalities, even when co-occurring, are not always tied to a unitary neurofunctional disturbance. Rather, it is possible that the right hemisphere assumes an integral (albeit currently ill-defined) role within a neural network dedicated to prosody (Blonder et al., 1989; Pell, 1996), and that expressive and receptive difficulties for prosody often stem from independently-compromised mechanisms in RHD patients. Certainly, indications in Experiment 1 that the RHD patients' expressive deficits were mostly *quantitative* in nature are inconsistent with the notion of a central deficit, or loss of knowledge about prosodic representations affecting both expressive and receptive channels, in the current RHD sample (see also Gandour et al., 1995; Baum & Pell, 1997). Through increased knowledge of how the psycholinguistic processes underlying prosody *interact* in both speech production and perception, future undertakings will be better equipped to arrive at a neurofunctional description of prosody that more closely reflects the demands inherent in natural speech.

REFERENCES

Alexander, M. P. and Bakchine, S. (1994). "Reliability and interpretation of perceptual analyses of prosodic impairments." Unpublished manuscript.

Bachorowski, J. and Owren, M. J. (1995). "Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context." Psychological Science 6(4): 219-224.

Baum, S. R. (Under review). "The role of fundamental frequency and duration in the perception of linguistic stress."

Baum, S. R., Daniloff, J. K., Daniloff, R. and Lewis, J. (1982). "Sentence comprehension by Broca's aphasics: Effects of some suprasegmental variables." Brain and Language 17: 261-271.

Baum, S. R. and Pell, M. D. (1997). "Production of affective and linguistic prosody by brain-damaged patients." Aphasiology 11: 177-198.

Baum, S. R. and Pell, M. D. (In Press). The neural bases of prosody: Insights from lesion studies and neuroimaging. In M. Lynch (Ed.), The Cognitive Science of Prosody: Interdisciplinary Perspectives. Amsterdam: North Holland/Elsevier.

Beach, C. M. (1991). "The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations." Journal of Memory and Language 30: 644-663.

Beckman, M. E. (1996). "The parsing of prosody." Language and Cognitive Processes **11**(1/2): 17-67.

Beckman, M. E. and Pierrehumbert, J. B. (1986). "Intonational structure in Japanese and English." Phonology Yearbook **3**: 255-309.

Behrens, S. J. (1985). "The perception of stress and lateralization of prosody." Brain and Language **26**: 332-348.

Behrens, S. J. (1988). "The role of the right hemisphere in the production of linguistic stress." Brain and Language **33**: 104-127.

Behrens, S. J. (1989). "Characterizing sentence intonation in a right hemisphere-damaged population." Brain and Language **37**: 181-200.

Bever, T. G. (1981). Normal acquisition processes explain the critical period for language learning. In K. Diller (Ed.), Individual Differences and Universals in Language Learning Aptitude. Rowley, MA: Newbury House.

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

Blonder, L. X., Gur, R. E. and Gur, R. C. (1989). "The effects of right and left hemiparkinsonism on prosody." Brain and Language **36**: 193-207.

Blonder, L. X., Pickering, J. E., Heath, R. L., Smith, C. D. and Butler, S. M. (1995). "Prosodic characteristics of speech pre- and post-right hemisphere stroke." Brain and Language **51**(2): 318-335.

Blumstein, S. and Cooper, W. E. (1974). "Hemispheric processing of intonation contours." Cortex **10**: 146-158.

Blumstein, S. and Goodglass, H. (1972). "The perception of stress as a semantic cue in aphasia." Journal of Speech and Hearing Research **15**: 800-806.

Bolinger, D. L. (1978). Intonation across languages. In J. H. Greenberg (Ed.), Universals of Human Language. Stanford, California, Stanford University Press: 471-524.

Bolinger, D. L. (1955). "Intersections of stress and intonation." Word **11**(2): 195-203.

Bolinger, D. L. (1958). "A theory of pitch accent in English." Word **14**: 109-149.

Bolinger, D. L. (1961). "Contrastive accent and contrastive stress." Language, 37, 83-96.

Borod, J. C., Koff, E., Perlman Lorch, M. and Nicholas, M. (1985). "Channels of emotional expression in patients with unilateral brain damage." Archives of Neurology **42**: 345-348.

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E. and Diller, L. (1990). "Parameters of emotional processing in neuropsychiatric disorders: Conceptual issues and a battery of tests." Journal of Communication Disorders **23**: 247-271.

Bowers, D., Coslett, H. B., Bauer, R. M., Speedie, L. J. and Heilman, K. M. (1987). "Comprehension of emotional prosody following unilateral hemispheric lesions: Processing defect versus distraction defect." Neuropsychologia **25**(2): 317-328.

Brådvik, B., Dravins, C., Holtås, S., Rosén, I, Ryding, E. and Ingvar, D. H. (1990). "Do single right hemisphere infarcts or transient ischaemic attacks result in aprosody?" Acta Neurologica Scandanavica **81**: 61-70.

Brådvik, B., Dravins, C., Holtås, S., Rosén, I, Ryding, E. and Ingvar, D. H. (1991). "Disturbances of speech prosody following right hemisphere infarcts." Acta Neurologica Scandanivica **54**(2): 114-126.

Brown, B. L., Strong, W. J. and Rencher, A. C. (1974). "Fifty-four voices from two: The effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech." Journal of the Acoustical Society of America **55**(2): 313-318.

Brown, W.S. and McGlone, R. E. (1974). "Aerodynamic and acoustic study of stress in sentence productions." Journal of the Acoustical Society of America **56**(3): 971-974.

Bryan, K. L. (1989). "Language prosody and the right hemisphere." Aphasiology **3**(4): 285-299.

Cancelliere, A. and Hausdorf, P. (1988). "Emotional expression in Huntington's disease." Journal of Clinical and Experimental Neuropsychology **10**: 62.

Cancelliere, A. and Kertesz, A. (1990). "Lesion localization in acquired deficits of emotional expression and comprehension." Brain and Cognition **13**: 133-147.

Caplan, D. (1992). Language: Structure, process, and disorders. Cambridge, MIT Press.

Carmon, A. and Nachshon, I. (1971). "Effect of unilateral brain damage on perception of temporal order." Cortex **7**: 410-418.

Cohen, A., Collier, R. and 't Hart, J. (1982). "Declination: Construct or intrinsic feature of speech pitch?" Phonetica **39**: 254-273.

Cohen, A. and 't Hart, J. (1967). "On the anatomy of intonation." Lingua **19**: 177-192.

Colsher, P. L., Cooper, W. E. and Graff-Radford, N. (1987). "Intonational variability in the speech of right-hemisphere damaged patients." Brain and Language **32**: 379-383.

Cooper, W. E., Eady, S. J. and Mueller, P. R. (1985). "Acoustical aspects of contrastive stress in question-answer contexts." Journal of the Acoustical Society of America 77(6): 2142-2156.

Cooper, W. E., Soares, C., Ham, A and Damon, K. (1983). "The influence of inter- and intra-speaker tempo on fundamental frequency and palatalization." Journal of the Acoustical Society of America **73**(5): 1723-1730.

Cooper, W. E., Soares, C., Nicol, J., Michelow, D. and Goloskie, S. (1984). "Clausal intonation after unilateral brain damage." Language and Speech **27**(1): 17-24.

Cosmides, L. (1983). "Invariances in the acoustic expression of emotion during speech." Journal of Experimental Psychology: Human Perception and Performance **9**(6): 864-881.

Cummings, K. E. and Clements, M. A. (1995). "Analysis of the glottal excitation of emotionally styled and stressed speech." Journal of the Acoustical Society of America **98**(1): 88-98.

Danly, M., Cooper, W. E. and Shapiro, B. (1983). "Fundamental frequency, language processing, and linguistic structure in Wernicke's aphasia." Brain and Language **19**: 1-24.

Danly, M. and Shapiro, B. (1982). "Speech prosody in Broca's aphasia." Brain and Language **16**: 171-190.

Darby, D. G. (1993). "Sensory aprosodia: A clinical clue to lesions of the inferior division of the right middle cerebral artery?" Neurology **43**: 567-572.

Darley, F. L., Aronson, A. E. and Brown, J. R. (1969). "Differential diagnostic patterns of dysarthria." Journal of Speech and Hearing Research **12**: 246-269.

Davitz, J. R. (1964). Auditory correlates of vocal expressions of emotional meanings. In J. R. Davitz (Ed.), The Communication of Emotional Meaning. New York, McGraw-Hill: 101-112.

Davitz, J. R. (1964). Personality, perceptual, and cognitive correlates of emotional sensitivity. In J. R. Davitz (Ed.), The Communication of Emotional Meaning. New York, McGraw-Hill: 57-68.

Denes, P. (1959). "A preliminary investigation of certain aspects of intonation." Language and Speech **2**(2): 107-122.

Denes, P. and Milton-Williams, J. (1962). "Further studies in intonation." Language and Speech **5**(1): 1-14.

Dykstra, K., Gandour, J. and Stark, R. E. (1995). "Disruption of prosody after frontal lobe seizures in the non-dominant hemisphere." Aphasiology **9**(5): 453-476.

Eady, S. J. and Cooper, W. E. (1986). "Speech intonation and focus location in matched statements and questions." Journal of the Acoustical Society of America **80**(2): 402-415.

Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R. and Lotts, D. W. (1986). "Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments." Language and Speech **29**(3): 233-251.

Edmondson, J. A., Chan, J., Seibert, G. B. and Ross, E. D. (1987). "The effect of right-brain damage on acoustical measures of affective prosody in Taiwanese patients." Journal of Phonetics **15**: 219-233.

Eefting, W. (1990). "The effect of "information value" and "accentuation" on the duation of Dutch words, syllables, and segments." Journal of the Acoustical Society of America **89**(1): 412-424.

Ehlers, L. and Dalby, M. (1987). "Appreciation of emotional expressions in the visual and auditory modality in normal and brain-damaged patients." Acta Neurologica Scandinavica **76**: 251-256.

Ekman, P., Levenson, R. W. and Friesen, W. V. (1983). "Autonomic nervous system activity distinguishes among emotions." Science **221**: 1208-1210.

Emmorey, K. D. (1987). "The neurological substrates for prosodic aspects of speech." Brain and Language **30**: 305-320.

Fairbanks, G. and Pronovost, W. (1939). "An experimental study of the pitch characteristics of the voice during the expression of emotion." Speech Monographs **6**: 87-104.

Farnetani, E., Taylor Torsello, C. and Cosi, P. (1988). "English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study." Language and Speech **31**(2): 157-180.

Fernald, A. (1989). "Intonation and communicative intent in mothers' speech to infants: Is the melody the message?" Child Development **60**: 1497-1510.

Fernald, A. and Simon, T. (1984). "Expanded intonation contours in mothers' speech to newborns." Developmental Psychology **20**(1): 104-113.

Ferreira, F. (1993). "Creation of prosody during sentence production." Psychological Review **100**(2): 233-253.

Fonagy, I. (1978). "A new method of investigating the perception of prosodic features." Language and Speech **21**: 34-49.

Frick, R. W. (1985). "Communicating emotion: The role of prosodic features." Psychological Bulletin **97**(3): 412-429.

Fry, D. B. (1955). "Duration and intensity as physical correlates of linguistic stress." Journal of the Acoustical Society of America **27**(4): 765-768.

Fry, D. B. (1958). "Experiments in the perception of stress." Language and Speech **1**: 126-152.

Fry, D. B. (1970). Prosodic phenomena. In B. Malmberg (Ed.), Manual of Phonetics. Amsterdam, North Holland: 365-410.

Gandour, J. and Dardarananda, R. (1983). "Identification of tonal contrasts in Thai aphasic patients." Brain and Language **18**: 98-114.

Gandour, J., Holasuit Petty, S. and Dardarananda, R. (1989). "Dysprosody in Broca's aphasia: A case study." Brain and Language **37**: 232-257.

Gandour, J., Larsen, J., Dechongkit, S., Ponglorpisit, S. and Khunadorn, F. (1995). "Speech prosody in affective contexts in Thai patients with right hemisphere lesions." Brain and Language **51**(3): 422-443.

Gandour, J., Ponglorpisit, S., Khunadorn, F., Dechongkit, S., Boongird, P., Boonklam, R. and Potisuk, S. (1992). "Lexical tones in Thai after unilateral brain damage." Brain and Language **43**: 275-307.

Gauthier, L., Dehaut, F. and Joanette, Y. (1989). "The Bells Test: A quantitative and qualitative test for visual neglect." International Journal of Clinical Neuropsychology **11**(2): 49-54.

Gorelick, P. B. and Ross, E. D. (1987). "The aprosodias: Further functional-anatomic evidence for the organization of affective language in the right hemisphere." Journal of Neurology, Neurosurgery, and Psychiatry **50**: 553-560.

Hadding-Koch, K. and Studdert-Kennedy, M. (1964). "An experimental study of some intonation contours." Phonetica **11**: 175-185.

Haggard, M., Ambler, S. and Callow, M. (1970). "Pitch as a voicing cue." Journal of the Acoustical Society of America **47**: 613-617.

Halle, P. A. and Vihman, M. M. (1991). "Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants." Language and Speech **34**(4): 299-318.

Hauser, M. D. and Fowler, C. A. (1992). "Fundamental frequency declination is not unique to human speech: Evidence from nonhuman primates." Journal of the Acoustical Society of America **91**(1): 363-369.

Heilman, K. M., Bowers, D., Speedie, L. and Coslett, H. B. (1984). "Comprehension of affective and nonaffective prosody." Neurology **34**: 917-921.

Heilman, K. M., Scholes, R. and Watson, R. T. (1975). "Auditory affective agnosia: Disturbed comprehension of affective speech." Journal of Neurology, Neurosurgery, and Psychiatry **38**: 69-72.

Hird, K. and Kirsner, K. (1993). "Dysprosody following acquired neurogenic impairment." Brain and Language **45**(1): 46-60.

Hughes, C. P., Chan, J. L. and Su, M. S. (1983). "Aprosodia in Chinese patients with right cerebral hemisphere lesions." Archives of Neurology **40**: 732-737.

Huttar, G. L. (1968). "Relations between prosodic variables and emotions in normal American English utterances." Journal of Speech and Hearing Research **11**: 467-480.

Johnson, W. F., Emde, R. N., Scherer, K. R. and Linnert, M. D. (1986). "Recognition of emotion from vocal cues." Archives of General Psychiatry **43**: 280-283.

Kaplan, E. L. and Kaplan, G. (1971). The prelinguistic child. In J. Elliot (Ed.), Human Development and Cognitive Processes. New York, Holt, Rinehart, and Winston: 358-381.

Kaplan, J. A., Brownell, H., Jacobs, J. R. and Gardner, H. (1990). "The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks." Brain and Language **38**: 315-333.

Kent, R. D. and Rosenbek, J. C. (1982). "Prosodic disturbance and neurologic lesion." Brain and Language **15**: 259-291.

Kimelman, M. (1991). "The role of target word stress in auditory comprehension of aphasic listeners." Journal of Speech and Hearing Research **34**: 334-339.

Kimelman, M. and McNeil, M. (1987). "An investigation of emphatic stress comprehension in adult aphasia: A replication." Journal of Speech and Hearing Research **30**: 295-300.

Kimelman, M. and McNeil, M. (1989). Contextual influences on the auditory comprehension of normally stressed targets b y aphasic listeners. In T. E. Prescott (Ed.), Clinical Aphasiology. Boston, MA, Little-Brown: 407-420.

Kimura, D. (1961). "Cerebral dominance and the perception of verbal stimuli." Canadian Journal of Psychology **15**: 166-171.

Klatt, D. H. (1976). "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence." Journal of the Acoustical Society of America **59**(5): 1208-1221.

Kramer, E. (1964). "Elimination of verbal cues in judgments of emotion from voice." Journal of Abnormal and Social Psychology **68**(4): 390-396.

Ladd, D. R., Silverman, K., Tolkmitt, F., Bergmann, G. and Scherer, K. R. (1985). "Evidence for the independent function of intonation contour type, voice quality, and Fo range in signaling speaker affect." Journal of the Acoustical Society of America **78**(2): 435-444.

Lea, W. A. (1977). Acoustic correlates of stress and juncture. In L. M. Hyman (Ed.), Studies in Stress and Accent. Los Angeles, Dept. of Linguistics, U.C.L.A.: 83-120.

Lebrun, Y., Lessinnes, A., De Vresse, L and Leleux, C. (1985). "Dysprosody and the non-dominant hemisphere." Language Sciences 7(1): 41-52.

Lehiste (1970). Suprasegmentals. Cambridge, MIT Press.

Lehiste, I., Olive, J. P. and Streeter, L. A. (1976). "Role of duration in disambiguating syntactically ambiguous sentences." Journal of the Acoustical Society of America **60**(5): 1199-1202.

Lenneberg, E. H. (1967). Biological Foundations of Language. New York, John Wiley & Sons, Inc.

Ley, R. G. and Bryden, M. P. (1982). "A dissociation of right and left hemispheric effects for recognizing emotional tone and verbal content." Brain and Cognition **1**: 3-9.

Lieberman, P. (1960). "Some acoustic correlates of word stress in American English." Journal of the Acoustical Society of America **32**(4): 451-454.

Lieberman, P. (1961). "Perturbations in vocal pitch." Journal of the Acoustical Society of America **33**(5): 597-603.

Lieberman, P. (1965). "On the acoustic basis of the perception of intonation by linguists." Word **21**: 40-54.

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R. and Miller, M. (1985). "Measures of the sentence intonation of read and spontaneous speech in American English." Journal of the Acoustical Society of America **77**(2): 649-657.

Lieberman, P. and Michaels, S. B. (1962). "Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech." Journal of the Acoustical Society of America **34**(7): 922-927.

Lyberg, B. (1979). "Final lengthening--partly a consequence of restrictions on the speech of fundamental frequency change?" Journal of Phonetics **7**: 187-196.

MacWhinney, B. and Bates, E. (1978). "Sentential devices for conveying givenness and newness: A cross-cultural developmental study." Journal of Verbal Learning and Verbal Behavior **17**: 539-558.

Majewski, W. and Blasdell, R. (1968). "Influence of fundamental frequency cues on the perception of some synthetic intonation contours." Journal of the Acoustical Society of America **45**(2): 450-457.

McClean, M. D. and Tiffany, W. R. (1973). "The acoustic parameters of stress in relation to syllable position, speech loudness and rate." Language and Speech **16**(3): 283-290.

McRoberts, G., Studdert-Kennedy, M. and Shankweiler, D. P. (1995). "The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation." Perception & Psychophysics **57**(2): 159-174.

Mertus, J. (1989). BLISS User's Manual. Providence, R.I., Brown University.

Monrad-Krohn, G. H. (1947). "Dysprosody or altered "melody of language"." Brain **70**: 405-423.

Morgan, J. L., Meier, R. P. and Newport, E. L. (1987). "Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language." Cognitive Psychology **19**: 498-550.

Morton, J. and Jassem, W. (1965). "Acoustic correlates of stress." Language and Speech **8**(3): 159-181.

Murray, I. R. and Arnott, J. L. (1993). "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion." Journal of the Acoustical Society of America **93**(2): 1097-1108.

Nooteboom, S. G. and Kruyt, J. G. (1987). "Accents, focus distribution, and the perceived distribution of given and new information: An experiment." Journal of the Acoustical Society of America **82**(5): 1512-1524.

Ohala, J. J. (1977). The physiology of stress. In S. Hyman (Ed.), Studies in Stress and Accent. California.

Ohala, J. J. (1984). "An ethological perspective on common cross-language utilization of F_0 of voice." Phonetica **41**: 1-16.

O'Shaughnessy, D. (1979). "Linguistic features in fundamental frequency patterns." Journal of Phonetics **7**: 119-145.

O'Shaughnessy, D. and Allen, J. (1983). "Linguistic modality effects on fundamental frequency in speech." Journal of the Acoustical Society of America **74**(4): 1155-1171.

Ouellette, G. P. and Baum, S. R. (1993). "Acoustic analysis of prosodic cues in left- and right-hemisphere-damaged patients." Aphasiology **8**(3): 257-283.

Packard, J. L. (1986). "Tone production deficits in nonfluent aphasic Chinese speech." Brain and Language **29**: 212-223.

Pakosz, M. (1983). "Attitudinal judgments in intonation: Some evidence for a theory." Journal of Psycholinguistic Research **12**(3): 311-326.

Pashek, G. V. and Brookshire, R. H. (1982). "Effects of rate of speech and linguistic stress on auditory paragraph comprehension of aphasic individuals." Journal of Speech and Hearing Research **25**: 377-383.

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

Pell, M. D. and S. R. Baum (1997a). "The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults." Brain and Language **57**(1): 80-99.

Pell, M. D. and S. R. Baum (1997b). "Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody." Brain and Language **57**(2): 195-214.

Pierrehumbert, J. (1979). "The perception of fundamental frequency declination." Journal of the Acoustical Society of America **66**(2): 363-369.

Pike, K. L. (1945). The Intonation of American English. Ann Arbor, MI, University of Michigan.

Pollack, I., Rubenstein, H. and Horowitz, A. (1960). "Communication of verbal modes of expression." Language and Speech **3**: 121-130.

Price, P. J., Ostendorf, S., Shattuck-Hufnagel, S. and Fong, C. (1991). "The use of prosody in syntactic disambiguation." Journal of the Acoustical Society of America **90**(6): 2956-2970.

Robin, D. A., Tranel, D. and Damasio, H. (1990). "Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions." Brain and Language **39**: 539-555.

Ross, E. D. (1981). "The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere." Archives of Neurology **38**: 561-569.

Ross, E. D. (1988). "Prosody and brain lateralization." Archives of Neurology **45**: 338-339.

Ross, E. D. (1993). "Nonverbal aspects of language." Behavioral Neurology **11**(1): 9-23.

Ross, E. D., Edmondson, J. A. and 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 **14**: 283-302.

Ross, E. D., Edmondson, J. A., Seibert, G. B. and 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 and Language **33**: 128-145.

Ross, E. D., Harney, J. H., deLacoste-Utamsing, C. and Purdy, P. D. (1981). "How the brain integrates affective and propositional language into a unified behavioral function." Archives of Neurology **38**: 745-748.

Ross, E. D. and Mesulam, M. (1979). "Dominant language functions of the right hemisphere?: Prosody and emotional gesturing." Archives of Neurology **36**: 144-149.

Ross, E. D., Thompson, R. D. and Yenkosky, J. (1997). "Lateralization of affective prosody in brain and the collosal integration of hemispheric language functions." Brain and Language **56**: 27-54.

Ryalls, J. (1986). "What constitutes a primary disturbance of speech prosody?: A reply to Shapiro and Danly." Brain and Language **29**: 183-187.

Ryalls, J. (1988). "Concerning right-hemisphere dominance for affective language." Archives of Neurology **45**: 337-338.

Ryalls, J., Joanette, Y. and Feldman, L. (1987). "An acoustic comparison of normal and right-hemisphere-damaged speech prosody." Cortex **23**: 685-694.

Ryalls, J., Le Dorze, G., Lever, N., Ouellet, L. and Larfeuil, C. (1994). "The effects of age and sex on speech intonation and duration for matched statements and questions in French." Journal of the Acoustical Society of America **95**(4): 2274-2276.

Ryalls, J. and Reinvang, I. (1986). "Functional lateralization of linguistic tones: Acoustic evidence from Norwegian." Language and Speech **29**(4): 389-398.

Scherer, K. R. (1974). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.), Non-verbal Communication. New York, Oxford University Press: 105-111.

Scherer, K. R. (1986). "Vocal affect expression: A review and a model for future research." Psychological Bulletin **99**(2): 143-165.

Schlanger, B. B., Schlanger, P. and Gerstman, L. J. (1976). "The perception of emotionally toned sentences by right hemisphere-damaged and aphasic subjects." Brain and Language **3**: 396-403.

Scott, S., Caird, F. and Williams, B. (1984). "Evidence for an apparent sensory speech disorder in Parkinson's disease." Journal of Neurology, Neurosurgery, and Psychiatry **47**: 840-843.

Shapiro, B. and Danly, M. (1985). "The role of the right hemisphere in the control of speech prosody in propositional and affective contexts." Brain and Language **25**: 19-36.

Sidtis, J. J. and Feldmann, E. (1990). "Transient ischemic attacks presenting with a loss of pitch perception." Cortex **26**: 469-471.

Siegman, A. W., Dembroski, T. M. and Crump, D. (1992). "Speech rate, loudness, and cardiovascular reactivity." Journal of Behavioral Medicine **15**(5): 519-532.

Speedie, L. J., Brake, N., Folstein, S. E., Bowers, D. and Heilman, K. M. (1990). "Comprehension of prosody in Huntington's disease." Journal of Neurology, Neurosurgery, and Psychiatry **53**: 607-610.

Starkstein, S. E., Federoff, J. P., Price, T. R., Leiguarda, R. C. and Robinson, R. G. (1994). "Neuropsychological and neuroradiologic correlates of emotional prosody comprehension." Neurology **44**: 515-522.

Streeter, L. (1978). "Acoustic determinants of phrase boundary perception." Journal of the Acoustical Society of America **64**: 1582-1592.

Studdert-Kennedy, M. and Hadding, K. (1973). "Auditory and linguistic processes in the perception of intonation contours." Language and Speech **16**: 293-313.

Swerts, M. and Geluykens, R. (1993). "The prosody of information units in spontaneous monologue." Phonetica **50**: 189-196.

Tompkins, C. A. and Flowers, C. R. (1985). "Perception of emotional intonation by brain-damaged adults: The influence of task processing levels." Journal of Speech and Hearing Research **28**: 527-538.

Tucker, D. M., Watson, R. T. and Heilman, K. M. (1977). "Discrimination and evocation of affectively intoned speech in patients with right parietal disease." Neurology **27**: 947-950.

Turk, A. E. and Sawusch, J. R. (1996). "The processing of duration and intensity cues to prominence." Journal of the Acoustical Society of America **99**(6): 3782-3790.

Uldall, E. (1960). "Attitudinal meanings conveyed by intonation contours." Language and Speech **3**(1): 223-234.

Umeda, N. (1982). ""F_0 declination" is situation dependent." Journal of Phonetics **10**: 279-290.

Van Lancker, D. (1980). "Cerebral lateralization of pitch cues in the linguistic signal." Papers in Linguistics **13**: 201-277.

Van Lancker, D. and Fromkin, V. (1973). "Hemispheric specialization for pitch and "tone": Evidence from Thai." Journal of Phonetics **1**: 101-109.

Van Lancker, D. and Pachana, N. (1995). "Acquired dysprosodic speech production: Mood, motivational, cognitive, or motor disorder?" Brain and Language **51**(1): 193-196.

Van Lancker, D. and 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**: 963-970.

Vande Kopple, W. J. (1982). "The given-new strategy of comprehension and some natural expository paragraphs." Journal of Psycholinguistic Research **11**(5): 501-520.

Weintraub, S., Mesulam, M. and Kramer, L. (1981). "Disturbances in prosody: A right-hemisphere contribution to language." Archives of Neurology **38**: 742-745.

Weismer, G. and Ingrisano, D. (1979). "Phrase-level timing patterns in English: Effects of emphatic stress location and speaking rate." Journal of Speech and Hearing Research **22**: 516-533.

Williams, C. E. and Stevens, K. N. (1972). "Emotions and speech: Some acoustical correlates." The Journal of the Acoustical Society of America **52**(2): 1238-1250.

Wingfield, A., Lombardi, L. and Sokol, S. (1984). "Prosodic features and the intelligibility of accelerated speech: Syntactic versus periodic segmentation." Journal of Speech and Hearing Research **27**: 128-134.

Zatorre, R. (1988). "Pitch perception of complex tones and human temporal-lobe function." Journal of the Acoustical Society of America **84**: 566-572.

Zatorre, R., Evans, A. and Meyer, E. (1994). "Neural mechanisms underlying melodic perception and memory for pitch." Journal of Neuroscience **14**: 1908-1919.

APPENDIX A

For both short and long stimuli, the following effects were noted for the keyword duration data:

KEYWORD

Neutral: [$F_{SHORT}(2,36)=129.68$, $p<.001$; $F_{LONG}(3,36)=78.31$, $p<.001$]
Sad: [$F_{SHORT}(2,36)=73.51$, $p<.001$; $F_{LONG}(3,36)=72.86$, $p<.001$]
Happy: [$F_{SHORT}(2,36)=86.37$, $p<.001$; $F_{LONG}(3,36)=60.10$, $p<.001$]
Angry: [$F_{SHORT}(2,36)=72.56$, $p<.001$; $F_{LONG}(3,36)=86.29$, $p<.001$]

FOCUS

Neutral: [$F_{SHORT}(2,36)=12.87$, $p<.001$; $F_{LONG}(2,36)=13.99$, $p<.001$]
Sad: [$F_{SHORT}(2,36)=6.03$, $p<.01$; $F_{LONG}(2,36)=3.37$, $p<.05$]
Happy: [$F_{SHORT}(2,36)=5.04$, $p<.05$; $F_{LONG}(2,36)=14.26$, $p<.001$]
Angry: [$F_{SHORT}(2,36)=3.14$, $p=.055$; $F_{LONG}(2,36)=4.87$, $p<.05$]

KEYWORD X FOCUS

Neutral: [$F_{SHORT}(4,72)=34.63$, $p<.001$; $F_{LONG}(6,108)=30.59$, $p<.001$]
Sad: [$F_{SHORT}(4,72)=25.97$, $p<.001$; $F_{LONG}(6,108)=17.06$, $p<.001$]
Happy: [$F_{SHORT}(4,72)=32.88$, $p<.001$; $F_{LONG}(6,108)=20.85$, $p<.001$]
Angry: [$F_{SHORT}(4,72)=30.85$, $p<.001$; $F_{LONG}(6,108)=25.79$, $p<.001$]

For only the short stimuli, the following three-way interaction was observed:

KEYWORD X MODALITY X FOCUS

Neutral: [$F(4,72)=6.61$, $p<.001$]
Sad: [$F(4,72)=2.82$, $p<.05$]
Happy: [$F(4,72)=2.62$, $p<.05$]
Angry: [$F(4,72)=9.42$, $p<.001$]

APPENDIX B

For both short and long stimuli, the following effects were noted for the keyword F_0 data:

MODALITY

Neutral: [$F_{SHORT}(1,18)=194.62$, $p<.001$; $F_{LONG}(1,18)=151.14$, $p<.001$]
Sad: [$F_{SHORT}(1,18)=104.98$, $p<.001$; $F_{LONG}(1,18)=153.34$, $p<.001$]
Happy: [$F_{SHORT}(1,18)=68.77$, $p<.001$; $F_{LONG}(1,18)=59.23$, $p<.001$]
Angry: [$F_{SHORT}(1,18)=73.23$, $p<.001$; $F_{LONG}(1,18)=126.12$, $p<.001$]

KEYWORD X MODALITY

Neutral: [$F_{SHORT}(3,54)=77.28$, $p<.001$; $F_{LONG}(4,72)=124.11$, $p<.001$]
Sad: [$F_{SHORT}(3,54)=57.33$, $p<.001$; $F_{LONG}(4,72)=78.48$, $p<.001$]
Happy: [$F_{SHORT}(3,54)=41.47$, $p<.001$; $F_{LONG}(4,72)=62.62$, $p<.001$]
Angry: [$F_{SHORT}(3,54)=67.95$, $p<.001$; $F_{LONG}(4,72)=51.36$, $p<.001$]

KEYWORD X FOCUS

Neutral: [$F_{SHORT}(6,108)=12.26$, $p<.001$; $F_{LONG}(8,144)=7.09$, $p<.001$]
Sad: [$F_{SHORT}(6,108)=4.05$, $p=.001$; $F_{LONG}(8,144)=3.18$, $p<.01$]
Happy: [$F_{SHORT}(6,108)=9.23$, $p<.001$; $F_{LONG}(8,144)=7.62$, $p<.001$]
Angry: [$F_{SHORT}(6,108)=7.37$, $p<.001$; $F_{LONG}(8,144)=3.26$, $p<.01$]

MODALITY X FOCUS

Neutral: [$F_{SHORT}(2,36)=26.66$, $p<.001$; $F_{LONG}(2,36)=12.58$, $p<.001$]
Sad: [$F_{SHORT}(2,36)=15.23$, $p<.001$; $F_{LONG}(2,36)=7.77$, $p<.01$]
Happy: [$F_{SHORT}(2,36)=11.41$, $p<.001$; $F_{LONG}(2,36)=10.65$, $p<.001$]
Angry: [$F_{SHORT}(2,36)=10.36$, $p<.001$; $F_{LONG}(2,36)=7.99$, $p=.001$]

KEYWORD X MODALITY X FOCUS

Neutral: [$F_{SHORT}(6,108)=19.76$, $p<.001$; $F_{LONG}(8,144)=13.29$, $p<.001$]
Sad: [$F_{SHORT}(6,108)=5.35$, $p<.001$; $F_{LONG}(8,144)=4.75$, $p<.001$]
Happy: [$F_{SHORT}(6,108)=11.60$, $p<.001$; $F_{LONG}(8,144)=5.48$, $p<.001$]
Angry: [$F_{SHORT}(6,108)=8.01$, $p<.001$; $F_{LONG}(8,144)=3.75$, $p=.001$]

APPENDIX C

The following effects, unreported in the text, were significant for both the short and long stimuli in the analysis of F_0 accents:

EMOTION
[$F_{SHORT}(3,54) = 14.56, p<.001$; $F_{LONG}(3,54) = 16.55, p<.001$]

KEYWORD
[$F_{SHORT}(1,18) = 63.40, p<.001$; $F_{LONG}(1,18) = 82.06, p<.001$]

EMOTION X KEYWORD
[$F_{SHORT}(3,54) = 9.25, p<.001$; $F_{LONG}(3,54) = 9.68, p<.001$]

MODALITY X KEYWORD
[$F_{SHORT}(1,18) = 7.58, p<.05$; $F_{LONG}(1,18) = 14.66, p=.001$]

FOCUS X KEYWORD
[$F_{SHORT}(2,36) = 61.65, p<.001$; $F_{LONG}(2,36) = 48.90, p<.001$]

EMOTION X MODALITY
[$F_{SHORT}(3,54) = 4.35, p<.01$; $F_{LONG}(3,54) = 6.45, p=.001$]

The following effects, unreported in the text, were significant for the short stimuli only in the analysis of F_0 accents:

FOCUS
[$F(2,36) = 5.42, p<.01$]

MODALITY X FOCUS
[$F(2,36) = 7.91, p=.001$]

The following effects, unreported in the text, were significant for the long stimuli only in the analysis of F_0 accents:

MODALITY
[$F(1,18)=6.03, p<.05$]

EMOTION X FOCUS
[$F(6,108)=2.68, p<.05$]

APPENDIX D

(i) D-Neutral task - FOCUS X MODALITY X EMOTION Interaction

For the perception of sentence-*initial* emphasis when duration cues had been "neutralized", accuracy was significantly lower for sad stimuli than for happy, angry, and neutral stimuli, and significantly lower for neutral stimuli than for happy stimuli, when declarative utterances were presented. When interrogative utterances were presented, sentence-initial focus was recognized significantly better in happy and angry stimuli than neutral stimuli, and significantly better in angry stimuli than sad stimuli.

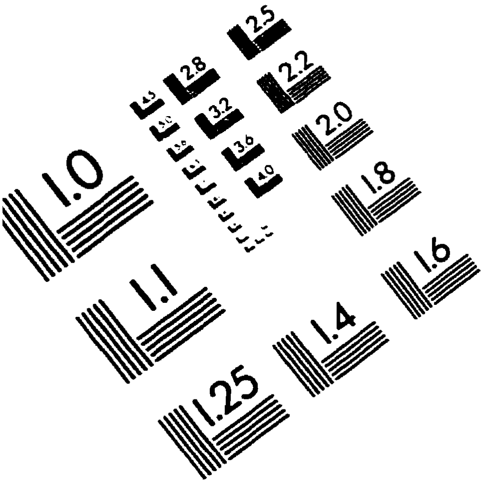
For the identification of sentence-*final* emphasis, accuracy was significantly lower for neutral stimuli than for happy and angry stimuli, and lower for sad stimuli than for angry stimuli, when declarative intonation was present. For interrogative utterances, terminal emphasis was perceived significantly better when presented in tandem with happy, angry, and neutral prosody than when presented with sad prosody.

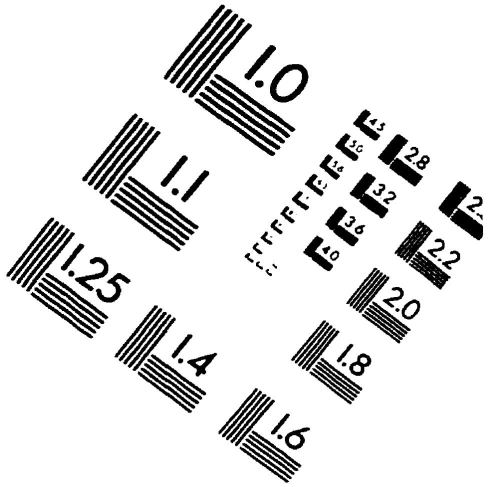
(ii) F-Neutral task- FOCUS X MODALITY X EMOTION Interaction

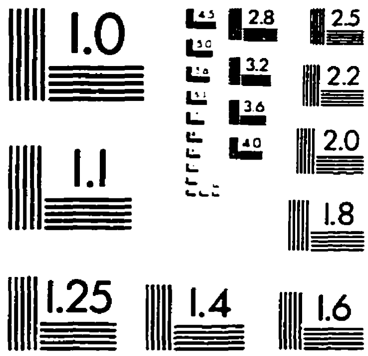
For sentence-*initial* focus, it was revealed that accuracy was significantly greater for angry and neutral stimuli than for sad stimuli, and significantly greater for angry stimuli than happy stimuli, when declaratives were presented. When interrogatives were presented, accuracy was significantly greater for happy and angry stimuli than for both sad and neutral stimuli.

For the perception of sentence-*final* focus, accuracy was greater for happy, angry, and sad stimuli relative to neutral stimuli, and for happy and angry stimuli relative to sad stimuli, when declaratives were presented. When interrogatives were presented, accuracy was significantly greater for happy, angry, and neutral stimuli than for sad stimuli.

IMAGE EVALUATION TEST TARGET (QA-3)

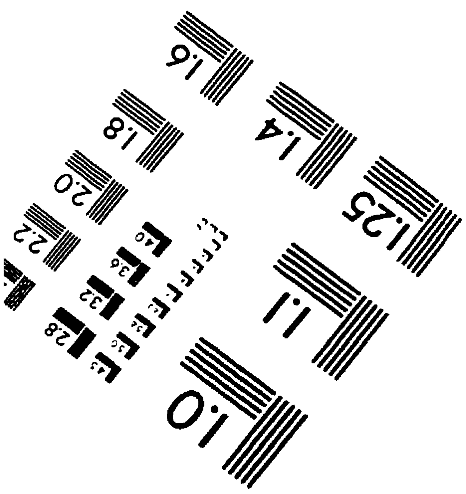


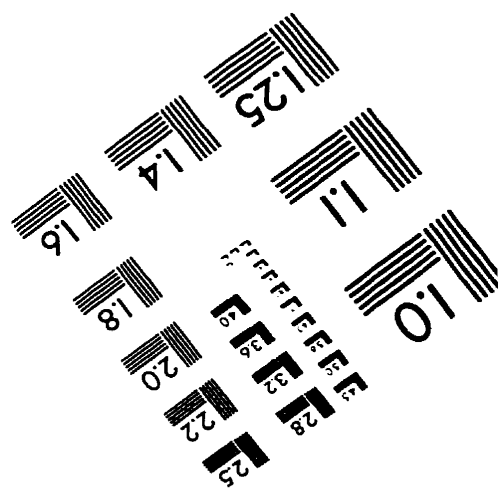




150mm

6"





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989