Selected acoustic characteristics of contrastive stress production in control geriatric, apraxic, and ataxic dysarthric speakers

JULIE M. LISS†
Department of Communication Disorders, 115 Shevlin Hall, University of Minnesota, Minneapolis, MN 55455, USA

GARY WEISMER
Department of Communicative Disorders and Waisman Center, University of Wisconsin–Madison, Madison, WI 53705, USA

(Received 8 February 1993; accepted 13 May 1993)

Abstract
Contrastive stress drills are often used in speech therapy to increase the intelligibility and communicative effectiveness of persons suffering from motor speech disorders. The rationale behind these drills is that the local effects of stress may improve articulatory performance on segments in the stressed word, as well as improve sentence-level prosodic adequacy. The purpose of the present investigation was to explore selected acoustic aspects of contrastive stress productions in control geriatrics and speakers with apraxia of speech and ataxic dysarthria. Results suggest that the phrase-level temporal and spectral effects of contrastive stress production among disordered speakers are not straightforward, and do not necessarily parallel those for normal speakers. These data are discussed relative to normal and disordered speech motor control.

Keywords: contrastive stress, motor speech disorders, acoustic analysis

Introduction
Among the variety of manipulations that have been purported to be effective in the treatment of motor speech disorders, contrastive stress drills are unique because they can be motivated by previous work in speech production (see Rosenbek and LaPointe, 1985). In the contrastive stress drill the patient is asked to produce heavy (emphatic) stress on one word of a multisyllabic utterance, usually in response to a question (see, for example, Yorkston, Beukelman, Minifie and Sapir, 1984). The motivation for this technique from the speech production literature includes the observation that stressed syllables usually are characterized by greater articulatory displacements (e.g. Kent and Netsell, 1971; Engstrand, 1988), and longer vocalic...
duration (Lehiste, 1970), when compared to unstressed syllables (although see Bolinger, 1961; Fry, 1955, p. 768; and Weismer and Ingrisano, 1979, Table 3, for exceptions to the classical view of increased duration as a correlate of stress). It has also been suggested that emphatic stress placed on a given syllable modifies the durations of the other syllables in the utterance, rather than restricting its influence to the local vicinity of the stressed syllable (Weismer and Ingrisano, 1979). The implication is that contrastive stress imparts both local and non-local changes in the timing profile of an utterance.

Presumably, increased articulatory displacements might increase the accuracy and precision of articulatory contacts in persons with motor speech disorders, many of whom have reduced ‘phonetic working spaces’ (see summary in Weismer and Martin, 1992). Moreover, variation in the overall timing profile of an utterance that results from a locally stressed syllable might be helpful in the normalization of dysprosodic utterances, another common characteristic of neurogenic speech disorders. Kearns and Simmons (1988, p. 608) have summarized this position by noting that the treatment utility of contrastive stress drills is to ‘direct attention to the durational and prosodic elements of the disorder’.

The purpose of the present study was to analyze in greater detail some acoustic characteristics of contrastive stress in neurologically normal speakers, speakers having apraxia of speech, and speakers having ataxic dysarthria. More specifically, we were interested in examining the phrase-level temporal effects of contrastive stress production, as well as the effects of contrastive stress on vocal tract gestures, as inferred from measurements of formant transitions. Apraxia of speech and ataxic dysarthria are motor speech disorders often characterized by disruptions of the temporal patterning of successive acoustic segments (Kent and Rosenbek, 1983; Kent, Netsell and Abbs, 1979). Moreover, apraxia of speech is often viewed as a disorder involving deficits in the positioning and smooth adjustment of articulatory structures (Kent and Rosenbek, 1983; Mcloch and Square, 1984). These considerations motivated the current focus on the speech timing and formant transition correlates of contrastive stress. Although contrastive stress is commonly discussed in theoretical and clinical writings on motor speech disorders, there are few relevant data in the literature. We therefore pursued this analysis to understand better the phenomenology of contrastively stressed utterances as it applies to speakers with normal and impaired speech motor control.

Methods

Subjects

Subjects included four each of control (C), apraxic (A), and ataxic dysarthric (D) speakers. All but two of the subjects (C1 and C2) were part of a larger investigation of neurogenic speech disorders and speech production, and detailed subject descriptions can be found in other published works (e.g. McNeil, Weismer, Adams and Mulligan, 1990; McNeil, Liss, Tseng and Kent, 1990). Briefly, all subjects were native speakers of American English and had speech discrimination scores of 70% or better at 40 dB hearing level in at least one ear. The four male apraxic subjects (59, 62, 54 and 72 years of age) had either left anterior or posterior lesions (as demonstrated by CAT scans), and were free of significant concomitant dysarthria or aphasia, as determined by performance on a battery of tests (including a structur-
al–functional examination, Raven Coloured Progressive Matrices, Word Fluency Measure, Porch Index of Communicative Ability, Revised Token Test, the Boston Diagnostic Aphasia Examination, and the Apraxia Battery for Adults). All four apraxic subjects had prominent speech characteristics traditionally associated with apraxia of speech (such as difficulty initiating speech, slow speaking rate, dysfluencies and articulatory errors). An impressionistic evaluation of the speech disorders among these patients is that one (A4) had a very severe apraxia of speech, whereas the other three apraxic patients could be labelled as moderate-to-severe. The two male and two female ataxic dysarthric subjects (44, 53, 55, and 32 years of age) were selected based on behavioural and imaging evidence of cerebellar lesion, and were judged to be free of cognitive and linguistic deficits by the same battery of tests listed above. Each of these latter subjects also had speech characteristics, as judged by the investigators, that suggested the presence of ataxic dysarthria. The impressionistic evaluation of speech disorder severity among these patients was that one subject was mild-to-moderate with the remaining three subjects in the moderate range. The four geriatric men who served as control subjects ranged in age from 63 to 69 years, and were judged to have normal speech and language function.

**Speech sample**

The speech sample of the larger investigation included a variety of single words, phrases and sentences that were elicited in several different speaking conditions. In the current study, productions of two sentences, 'buy Bobby a poppy', and 'build a big building', were analysed. These utterances were selected because they each contain three content words that can be contrastively stressed (buy, bobby, poppy; build, big, building), and each contain phoneme sequences (/ai/ from 'buy' and /tl/ from 'build') that are associated with relatively large and complex changes in articulatory configuration.

**Procedures**

The acoustic data for the present study were collected in the context of the larger investigation of electromyographic, kinematic, and aerodynamic analyses of neurogenic speech disorders (see Kent and McNeil, 1987, for a description of data collection methods). Whereas this report concerns only the acoustic measures, it is important to recognize that subjects (except C1 and C2) had electrodes in the labial musculature, movement transducers on the lips and jaw, and a pressure tube inside the mouth during the recording of the acoustic signal.

Subjects were asked to produce phrases and sentences following tape-recorded stimuli. Two speaking conditions were utilized in this study: neutral and contrastive stress. The two utterances analysed in this investigation were each produced five times in each condition, and these productions were randomly distributed among other sentences and phrases. Each sentence was produced five times in a neutral condition (for a total of 10 neutral tokens per subject), for which an acoustic model was provided but no specific directions about stress placement were given, and in a condition in which each of the content words was contrastively stressed in response to a tape-recorded question (five productions each in which the words **buy, bobby, poppy, build, big or building**, were contrastively stressed, for a total of 30 stressed...
tokens per subject). For example, the stimulus ‘should you give Bobby a poppy?’ was used to elicit the contrastively stressed response, ‘buy Bobby a poppy’. In addition to the auditory elicitation of these utterances, subjects were provided visual cues in the form of cards displaying the utterances with the stressed word written in capital letters.† In the ‘neutral’ condition, the subjects were asked to repeat the utterance provided by an auditory model. These tape-recorded models were produced as simple declarative utterances, and thus contained no contrastive stress. These utterances served as the ‘neutral stress’ exemplars (see Weismer and Ingrisano, 1979). That is, the acoustic measurements made on these utterances were used as ‘neutral’ reference points for the contrastive stress data.

**Data analysis**

**Perceptual analysis**

Four judges independently listened to all of the neutral and contrastive stress utterances. The perceptual analysis involved a determination of the lexical accuracy of each utterance, and in the contrastive stress conditions a judgement of stress placement accuracy. This step was taken to identify tokens that could not be used in all phases of the acoustic analysis, such as tokens that contained inaccurate stress placement in the contrastive stress condition, or those that contained more or fewer than four words. Three utterances were judged to have inaccurate or questionable stress placement and these were omitted from all analyses. Twenty-six utterances were identified that contained too many or too few lexical items (word additions or omissions). These tokens, 20 of which were produced by subject A2 (see footnote) were not included in the duration measurement analysis. However, 25 of these utterances were retained for the formant analysis of the /aI/ and /I/ segments because the lexical omissions or additions occurred after ‘buy Bobby’ or ‘build a big’.

**Acoustic analysis**

The acoustic analysis consisted of measurement of segment and utterance durations, based on conventional criteria found in the literature (see Klatt, 1976; Peterson and Lehiste, 1960; Weismer, 1984), and measurement of formant transition characteristics. The total utterance and segment durations were measured using the waveform editor facility of CSpeech, a multi-purpose speech analysis program (Milenkovic, 1990). Total utterance durations were measured as an index of global speaking rate. Because local duration effects (at the segmental level) were of interest in relating contrastively stressed to neutral-stress utterances, the information on global speaking rates could be of importance in the interpretation of segmental effects. Segment durations, consisting in the present case of vocalic intervals, were measured to reveal local and non-local effects of contrastive stress.

†Each subject was trained to follow this procedure until performance indicated that the task was understood. One subject (A2) had substantial difficulty performing the contrastive stress task and did not satisfactorily produce the first 15 utterances. These faulty trials were repeated at the end of the contrastive stress task until a full set of data was obtained. Total utterance duration measures were not conducted on many of A2’s productions because they contained an inextinguishable lexical addition (‘buy Bobby a big poppy’), however, formant measures for the /aI/ in ‘buy’ were performed in these cases.
Total utterance durations (TUD) were obtained by measuring the interval between the first glottal pulse of the first word in the utterance and the last clearly identifiable glottal pulse of the final word (for a description of relative timing patterns in some of these same subjects, see McNeil et al. 1990). The specific criteria for segment duration boundaries studied in the present investigation are provided in the Appendix. The vocalic durations chosen for this analysis were from the /aI/, /a/, and /a/ segments in /barba bjapa pi/, and the /i/, /I/, and /I/ segments in /bulabigrbuldn/. These particular vocalic segments were chosen because they were among the potentially stressed syllables in the utterances. The selection of vocalic segments in general (i.e. to the exclusion of consonantal segments) was based on two considerations. First, it is generally believed that the experimental manipulations that affect speech segment durations have greater effects on vowels, as compared to consonants (Gay, 1981). Second, the occurrence of syllable segregation (the separation of syllables in continuous speech by pauses or hesitations—see Kent and Rosenbek, 1983) in many of the current apraxic utterances made it impractical to perform reliable duration measures of consonant durations because of ambiguity in delineating true stop consonant boundaries from onsets and offsets of pause intervals. Whereas the effects on consonant intervals of emphatic stress may be interesting, they were not pursued further here because of these considerations.

Formant transition characteristics for /aI/ (‘buy’) and /I/ (‘build’) were measured with a Kay DSP 5500 workstation using a wide-band (300 Hz) spectrographic display and the associated waveforms, a scale expansion of 0–4 kHz and a screen time axis of 2 s. Transition characteristics are useful to draw inferences concerning the amount of change in vocal tract configuration (transition extent, or TE), the time over which that change occurs (transition duration, or TD), and the derived speed of that change (TE/TD = transition slope: see Weismer, Kent, Hodge and Martin, 1988). Previous investigations that quantified the transitional characteristics of formant trajectories have typically utilized either LPC values or time-frequency values obtained from digitized tracings of formant trajectories to detect changes of interest (see Weismer et al., 1988). In our early work with these particular speech samples, LPC errors were frequently detected due to the noisy source spectra (poor vocal quality) and occasional hypernasality of the disordered speakers in this study. Therefore an alternative method consisting of visual determination of transition segments was adopted. Quantitative transitional measures for second formant (F2) trajectories for the segments /aI/ in the word ‘buy’ and /I/ in the word ‘build’ were obtained from on-screen measurements using the Kay 5500 workstation and its spectrographic display function. This was accomplished by determining visually the steepest portion of the F2 trajectory (rising trajectory in the case of /aI/ and falling in the case of /I/)), setting vertical time cursors to identify the time segment over which the steepest portion of the trajectory occurred (transition duration, or TD), and setting horizontal frequency cursors at the centre of the formant band at each of these two time points (transition extent, or TE). Figure 1 shows an example of cursor placement for the /aI/ trajectory. The change in frequency (TE) was then divided by the change in time (TD) to determine the slope value for that portion of the trajectory. The focus on F2 transitions (to the exclusion of F1 or F3) for these quantitative analyses is motivated by the demonstrated sensitivity of this formant to the severity and phenomenology of motor speech disorders (Kent et al., 1989; Weismer, Martin, Kent and Kent, 1992).

Formant transition characteristics also were plotted as functions that relate the
transition extents to the transition durations, or ‘transition duration–extent functions’. It has been previously demonstrated for normal speakers and selected speakers with Parkinson’s disease and amyotrophic lateral sclerosis (Weismer, Kimelman and Gorman, 1985; Weismer et al., 1992; Weismer, 1992) that there is a lawful relationship between these measures, such that greater transition extents are typically associated with longer transition durations. This comparison was made with the present data by constructing such functions across stress conditions (i.e. where the transition duration was expected to vary by virtue of the induced durational effects), but within speaker groups, and then comparing the obtained functions across groups.

In addition to the quantitative assessment of transition characteristics, graphic analysis of the effects of contrastive stress on vocalic nuclei was accomplished by tracing whole first and second formant trajectories from hard copy, and then entering these traces into computer files via a graphics tablet. The trajectories were then plotted in superimposition to demonstrate the effects of interest. The combination of quantitative and qualitative analyses of formant trajectories is desirable because either type of analysis alone may miss important aspects of the speech production deficit in motor speech disorders (Weismer and Liss, 1991; Liss and Weismer, 1992). This technique also accommodated the small proportion of formants which were excluded from quantitative assessment because they could not be reliably measured. These formants typically corresponded with aberrant productions of the vowel which did not yield the expected trajectories for the segments of interest (/aI/ or /i/).
Contrastive stress

Reliability
A quasi-random subset of formants were remeasured by each of the three judges who performed the original measures. The vast majority of differences between the original and reliability slope calculations (TE/TD) were relatively small in magnitude. Approximately 90% of the intra-rater differences were less than 1.5 Hz/ms. Seventy-eight per cent of the inter-rater differences were less than 1.5 Hz/ms. Large discrepancies between original and reliability measures were reviewed and either corrected, as in the case of obvious measurement or recording error, or omitted from the analysis.

Results

Durations

Total utterance durations
Data on total utterance duration (TUD) are presented in Figure 2 for each subject in the three groups. For each subject the height of the four bars indicates the mean TUD for each of the four stress conditions in the two utterances; the error bars indicate one standard deviation around the mean. Data for A2's productions of 'Buy Bobby a poppy' are not included in this display (see footnote). The TUD data are presented mainly as indices of speaking rate differences across the groups. The shortest utterances were produced by normal speakers, the longest utterances by apraxic speakers, and the dysarthric TUDs fell roughly between these two extremes. The very long TUDs of some apraxic speakers (A1 and A4) are due in part to their production of relatively long pauses between syllables in the utterances.

The mean TUDs of the apraxic and dysarthric speakers are associated with greater variability than the mean TUDs of normal speakers, as judged by the length of the respective error bars. This may reflect in part the group differences in mean TUD magnitude. With regard to stress condition, the four control speakers exhibited similar patterns of total utterance duration for both utterances in which NEUTRAL utterances generally were produced more rapidly than the contrastively stressed utterances. If there is a similar stress pattern effect among the disordered subjects, it is subtle, and not displayed by subjects D2, D3 in 'Buy Bobby a poppy', or by subjects A1, A3, D2, D3 in 'Build a big building', where the NEUTRAL utterances have longer TUDs than one or more of the stressed utterances.

Relative durations
The effects of contrastive stress on selected vocalic durations in both utterances are shown in Figures 3–5 in the form of 'relative duration' plots. In these graphs the mean duration values obtained in the NEUTRAL stress condition are used as reference durations against which the mean vowel durations obtained in contrastive stress conditions are compared. Thus the horizontal line labelled 'zero' on the y-axis represents the mean reference durations, and the bar heights above the line indicate mean vocalic durations in the contrastive stress conditions that are greater than the same vocalic durations in the neutral stress condition; bar heights below the reference line indicate the opposite case. The contrastive stress conditions are indicated by the
Figure 2. Mean total utterance durations (TUD) and standard deviations for all subjects in each NEUTRAL and BUY, BOBBY, POPPY and BUILD, BIG and BUILDING stress conditions. The top row shows the mean TUDs for all subjects and conditions for the utterance, "buy Bobby a poppy". The bottom row shows the same measures for the utterance, "build a big building". Contrastive stress condition is coded in the leftmost panel for each utterance.
Figure 3. Relative duration plots showing the effects of contrastive stress on vocalic segments produced by the control speakers (C1–C4). The top row shows difference values for the segments /æ/ (‘buy’), /æ/ (‘bobby’) and /æ/ (‘poppy’). The bottom row shows the same measures for the segments /ɪ/ in ‘build’, /ɪ/ in big, and /ɪ/ in ‘building’. The solid line at zero represents the neutral duration. See text for additional details.
Relative Segment Durations: Apraxic

Figure 4. Relative duration plots showing the effects of contrastive stress on vocalic segments produced by apraxic speakers. See caption for Figure 3 for additional details.
Relative Segment Durations: Dysarthric

Figure 5. Relative duration plots showing the effects of contrastive stress on vocalic segments produced by dysarthric speakers. The triangles represent durations that are identical to those of the neutral condition, and therefore do not deviate from the baseline. See caption for Figure 3 for additional details.
bar codes given in the upper left-hand panel of each figure. For example, in Figure 3 the relative durations across vowels when BUY or BUILD were emphasized can be determined by following the unfilled bar heights across vowels; alternatively, the relative duration of a given vocalic nucleus across stress condition can be obtained by noting the values associated with the cluster of the three bars above a given vowel. This format allows for the identification of 'local' durational effects of the stress conditions relative to neutral, where 'local' refers to those changes occurring on the syllables that were contrastively stressed (such as the /ai/ in 'BUY Bobby a poppy'). This format also permits the description of 'non-local' effects or those durational changes occurring on syllables that were not contrastively stressed (such as both /i/ segments in 'build a big building').

The analysis of local and non-local temporal effects reported below is entirely descriptive, and forgoes the use of inferential statistics. The decision not to use inferential statistics was largely based on an initial examination of the data that suggested tremendous inter-subject variability in the response to contrastive stress production. Group comparisons using inferential statistics would not be meaningful, and would have masked some of the interesting phenomena described below.

Local effects
The current data on local effects suggest that whereas contrastive stress is often associated with a durational increment of the vocalic nucleus when the first word of an utterance is stressed (BUY or BUILD), stress on other words may produce a negligible duration change or even a shorter duration relative to the neutral stress condition (see BOBBY, POPPY and BIG stress effects for C1; POPPY for C2; and BIG for C4). Similar patterns characterize the apraxic and dysarthric subjects (Figures 4 and 5), with the following three qualifications. First, two subjects produced a negligible duration increment when the first word was stressed (D3, BUY; A3, BUILD), a condition that is otherwise associated with relatively substantial (25 ms or greater) temporal increments across all subjects. Second, the negative local changes described above for some control subjects above were generally of rather small magnitude, but in two apraxic subjects were very large (e.g. BOBBY, BIG and BUILDING for A1; BIG for A2). Third, there is a general impression that the neurogenically impaired subjects are more likely to produce large local deviations from the neutral stress condition regardless of sign, as compared to the normal subjects. Note, for example, the large local effects for BUY (D1, D4, A1), BOBBY (D4, A1, A3), POPPY (D1), BUILD (D4, A4) and BIG (D4, A4).

Non-local effects
Whereas previous work (Weismer and Ingrisano, 1979) showed that a contrastively stressed word resulted in fairly consistent adjustments of the duration of non-stressed words relative to their durations in neutral stressed utterances, the present data show no such consistency. There are cases in the present data where speakers ‘offset’ the positive duration increment on a stressed word with negative changes on non-stressed words (e.g. C4 for BOBBY and BUILDING, C1 and C3 for BUILD—Figure 3), but there were many exceptions to this pattern. Non-local effects in control speakers can therefore be summarized as idiosyncratic, both within (across stress conditions and/or utterances) and across speakers. The disordered speakers also show idiosyncratic patterns of non-local effects, but in a sense this is not unexpected because of the typically great amount of inter-subject variability in speech production data.
obtained from neurogenically disordered speakers. When negative non-local effects did occur, however, they were often of much greater magnitude than those observed for the control subjects (e.g. A1 and A4 for BUY and BUILD, D2 for BIG and BUILDING, D4 for BUY, BOBBY, BUILD, BIG, and BUILDING).

Transition characteristics

Transition duration--extent functions

Figure 6 shows transition duration-extent functions for the F2 transitional segment of /au/, and the data points on which they were based. The linear functions in each group were statistically significant. In the data from apraxic subjects the \( y \)-intercept is larger and the slope smaller than in the functions derived from the normal and dysarthric data. Only about 20% of the variance in TE is predicted by TD in the apraxic group (Pearson \( r=0.45 \)), whereas the linear functions in the control and dysarthric groups account for 34% \( (r=0.58) \) and 45% \( (r=0.67) \) of the variance, respectively. With regard to stress effects, for all three subject groups a greater proportion of data points in the BUY condition fall above, as compared to below, the regression lines. This reflects the tendency for stressed vocalic nuclei to have a TE increment that is somewhat greater than would be expected solely on the basis of the stress-related increment in TD. The distribution of points above or below the regression line in the NEUTRAL and BOBBY conditions is more even, with a slight tendency for the points to fall below the line.

Slopes

The means and standard deviations of the slope values for /au/ and /\( l \)/ are presented in Table 1. Note that the /au/ values are derived from the rising portion of the F2 trajectories in 'buy', and the /\( l \)/ values from the falling portion of the F2 trajectories in 'build'. Three out of the four control and dysarthric speakers, and two of the apraxic speakers, exhibited steeper slopes for the /au/ in 'buy' in the BOBBY condition than in the BUY condition. By the same token all of the control speakers, two of

![Figure 6. Scatterplots of transition extent (Hz) as a function of transition duration (ms), plus linear regression functions, shown separately for each group of speakers. Within each speaker group, data plotted are individual tokens; the code for stress condition is indicated at the bottom of the figure.](image)
Table 1. The mean and (standard deviation) of slope values for the trajectories /ai/ and /il/ by subject and speaking condition. C, A, and D correspond with group means.

<table>
<thead>
<tr>
<th></th>
<th>NEUTRAL</th>
<th>BUY</th>
<th>Bobby</th>
<th>NEUTRAL</th>
<th>BUILD</th>
<th>BIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ai/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>5.38(0.75)</td>
<td>6.44(0.69)</td>
<td>7.13(0.55)</td>
<td>8.86(1.14)</td>
<td>8.62(0.98)</td>
<td>9.48(2.06)</td>
</tr>
<tr>
<td>C2</td>
<td>5.3(1.75)</td>
<td>4.1(0.9)</td>
<td>4.8(1.11)</td>
<td>7.3(0.7)</td>
<td>7.0(0.76)</td>
<td>8.0(74)</td>
</tr>
<tr>
<td>C3</td>
<td>5.36(2.21)</td>
<td>3.95(1.22)</td>
<td>2.81(0.82)</td>
<td>8.1(1.1)</td>
<td>7.69(2.05)</td>
<td>9.62(1.48)</td>
</tr>
<tr>
<td>C4</td>
<td>4.27(0.25)</td>
<td>3.95(1.17)</td>
<td>4.07(0.75)</td>
<td>5.92(0.39)</td>
<td>6.54(0.76)</td>
<td>7.84(1.79)</td>
</tr>
<tr>
<td>C</td>
<td>5.08(1.4)</td>
<td>4.62(1.4)</td>
<td>4.7(1.8)</td>
<td>7.55(1.3)</td>
<td>7.5(1.4)</td>
<td>8.7(1.7)</td>
</tr>
<tr>
<td>A1</td>
<td>6.27(2.28)</td>
<td>3.38(1.19)</td>
<td>3.19(1.55)</td>
<td>—</td>
<td>4.15(1.26)</td>
<td>5.23(0.81)</td>
</tr>
<tr>
<td>A2</td>
<td>4.2(0.79)</td>
<td>3.7(0.77)</td>
<td>3.47(0.42)</td>
<td>6.07(0.84)</td>
<td>5.08(0.76)</td>
<td>5.91(1.14)</td>
</tr>
<tr>
<td>A3</td>
<td>3.9(1.3)</td>
<td>4.99(0.67)</td>
<td>5.03(1.47)</td>
<td>5.79(0.87)</td>
<td>5.48(1.58)</td>
<td>4.22(1.74)</td>
</tr>
<tr>
<td>A4</td>
<td>4.65(1.18)</td>
<td>5.38(0.88)</td>
<td>6.47(0.91)</td>
<td>3.36(0.29)</td>
<td>—</td>
<td>4.94(1.22)</td>
</tr>
<tr>
<td>A</td>
<td>4.7(1.6)</td>
<td>4.4(1.2)</td>
<td>4.5(1.9)</td>
<td>5.00(1.7)</td>
<td>4.96(1.3)</td>
<td>5.12(1.7)</td>
</tr>
<tr>
<td>D1</td>
<td>7.10(1.11)</td>
<td>5.92(2.05)</td>
<td>6.09(1.39)</td>
<td>7.69(0.95)</td>
<td>11.07(2.81)</td>
<td>8.95(1.31)</td>
</tr>
<tr>
<td>D2</td>
<td>3.42(0.52)</td>
<td>3.18(0.75)</td>
<td>3.8(0.52)</td>
<td>5.98(0.69)</td>
<td>7.4(1.43)</td>
<td>6.63(1.35)</td>
</tr>
<tr>
<td>D3</td>
<td>3.26(0.87)</td>
<td>4.12(0.55)</td>
<td>4.29(1.09)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D4</td>
<td>4.38(0.68)</td>
<td>4.31(0.74)</td>
<td>5.94(1.14)</td>
<td>7.99(0.73)</td>
<td>7.63(0.69)</td>
<td>10.42(2.56)</td>
</tr>
<tr>
<td>D</td>
<td>4.54(1.75)</td>
<td>4.43(1.43)</td>
<td>4.99(1.4)</td>
<td>6.67(1.9)</td>
<td>8.38(2.7)</td>
<td>8.44(2.4)</td>
</tr>
</tbody>
</table>

Given the large degree of intra-group variability evident in each of the three subject groups, as indicated by the ranges of mean slope values, interpretation of group trends must be viewed with caution. However, inspection of the group means (corresponding with rows labelled C, A, and D in Table 1) reveals that the slope values and standard deviations for /ai/ are similar across groups. By contrast, the apraxic group produced substantially shallower slopes for the /il/ segments than the other two groups. Mean slopes of the control and dysarthric groups for /il/ segments were roughly comparable.

Trajectories and stress

The data shown in Figure 6 suggest that there was a tendency for /ai/ to have larger TEs when the vocalic nuclei were stressed, compared to either the neutral stress or other-stress conditions. The issue of what happens to the formant trajectories, and by inference the rate and extent of change in vocal tract configuration, when a vocalic nucleus is stressed, can be addressed more completely by examination of the complete trajectories.

The effect of contrastive stress on normal trajectories is exemplified in Figure 7, which shows the F2s for C4's productions of the /il/ vowel nucleus (from 'build') in three different stress conditions. In this and the following figures, each panel shows the F2 trajectories for five repetitions of a stress condition by a single speaker. The top panels show data from the neutral stress condition, the middle panels show data from the condition where the measured vowel nucleus was stressed, and the
Contrastive stress

Figure 7. F2 trajectories for the /d/ nucleus in the word 'build', produced by subject C4. The top panel shows trajectories from the neutral stress condition, the middle panel from the build condition, and the bottom panel from the big condition. Within each panel, trajectories from multiple repetitions are shown.

Bottom panels show data from utterances where the stress occurred on a word other than the measured vowel nucleus. Note in Figure 7 that the slopes of the steeply falling parts of the trajectories are quite similar across the conditions, but slightly steeper in the big condition. This is consistent with the observation of shorter TDs and thus greater slope values in the big (and Bobby) condition. The great majority of normal stress effects were consistent with this description, and have been documented previously (Weismer and Liss, 1991).

Figures 8, 9, and 10 show several examples of stress effects on formant trajectories produced by disordered speakers in the present study. In Figure 8 the F2 patterns resulting from D1's productions of /d/ show similar transition slopes across the
stress conditions, and the tendency for a somewhat larger TE when the vowel nucleus is stressed. This is generally consistent with the pattern described here for normal speakers, and would be predicted from previous kinematic studies that show larger articulatory movements in stressed, as compared to unstressed vowel nuclei (Kent and Netsell, 1971; Ostry and Munhall, 1985). However, the current data from disordered speakers were characterized more by idiosyncratic stress effects, as shown in Figures 9 and 10. Figure 9 shows F1 and F2 trajectories from D4, who produced slightly greater F2 TEs for the diphthong /ai/ in the BUILD condition, but only at the expense of a shallower slope (compare the top to the middle panel; in fact, the steepest slopes were produced in the BOBBY condition). Figure 10 presents the most radical kind of stress effect observed in these data. This apraxic speaker (A4)
Figure 9. *F*1 and *F*2 trajectories for the /a/ nucleus in the word 'buy', produced by subject D4. The top panel shows trajectories from the NEUTRAL stress condition, the middle panel from the BUY condition, and the bottom panel from the BOBBY condition.

produced *F*2s for /a/ in the NEUTRAL condition that were clearly aberrant, but evidenced a truncated form of the expected falling trajectory. When the vocalic nucleus was stressed, the trajectories became quite disorganized (Figure 10, middle panel); when the utterance contained a contrastively stressed word (BIG) that did not involve the vowel nucleus of interest, three of the trajectories conformed to the expected normal pattern, whereas the remaining two were distinctly aberrant.

**Discussion**

This investigation examined selected phrase-level temporal and spectral effects of contrastive stress production among normal geriatric speakers, and individuals with
apraxia of speech and ataxic dysarthria. Disordered speakers showed less systematic local temporal effects than control speakers, and the effects that were seen tended to be of larger magnitude than those observed for the controls. Because non-local temporal effects were idiosyncratic even in the control speakers, the data from disordered speakers do not seem to provide any rich insight to the issue of phrase-level temporal patterning in these neurogenic speech disorders. However, many of the non-local effects that were observed in the apraxic and dysarthric speakers were often of greater magnitude than those seen for control speakers, paralleling the finding for local effects. One potential explanation for the lack of systematic non-local effects in the present control subjects is that older speakers, such as the ones used in the present investigation, fail to produce the phrase-level timing adjustments previously observed.

Figure 10. *F2* trajectories for the /l/ nucleus in the word 'build', produced by subject A4. See Figure 7 caption for additional details.
Contrastive stress

by Weismer and Ingrisano (1979) in young adult speakers. This ‘ageing explanation’, however, can probably be rejected, because Cariski (1992) has recently shown a lack of consistent non-local effects in both young adult and geriatric speakers. It may be that non-local timing effects associated with emphatic stress are highly idiosyncratic, and therefore not of much use as a probe of speech timing disruption in neurogenically disordered speakers.

The findings on formant trajectories indicated that these apraxic speakers have a tendency to produce larger transitions than the neurologically normal and dysarthric speakers, and that they adjust transition extent with changes in transition duration in a less orderly way than the speakers in the other two groups. Qualitative analysis also suggests that the effects of stress on formant trajectories are subtle in control speakers, and highly variable in disordered speakers. Similar observations have been made by Weismer and Liss (1991).

Implications for speech motor control in neurogenically disordered speakers

Perhaps the most interesting aspect of the speech motor control implications of the present data concerns the scaling of articulatory gestures. Many speech production theorists (Saltzman and Munhall, 1989; Kent, 1986; Fujimura, 1986) have been concerned about this issue, but it has not been discussed much in the literature on neurogenic speech disorders. In the present data both apraxic and dysarthric speakers produced many examples of what might be termed ‘overscaled’ articulatory gestures, as inferred from relatively large transition extents and segment durations. However, although both apraxic and dysarthric speakers produced some form of overscaling, the phenomenon appeared to be restricted to temporal events in dysarthrics (as revealed by excessive durational increments associated with local and non-local effects of stress), but characterized by both temporal and spectral events in apraxic speakers (as revealed by excessive temporal phenomena and an overall tendency for greater transition extents, indexed by the high y-intercept value in Figure 6). It would be of great interest to study the abilities of dysarthric and apraxic speakers to voluntarily grade articulatory displacements. If the scaling problem in apraxia is different from that in dysarthria, the present data would predict that the dysarthrics would be more successful than apraxics in producing graded articulatory displacements for a given vocalic nucleus (such as a diphthong). Studies such as this could help resolve long-standing controversies concerning the similarities and differences between dysarthric and apraxic speech symptomatology.

Clinical implications

Two common features of the apraxic utterances in this study were increased segment and total utterance durations, and generally greater TEs across all conditions as compared to the other subject groups. The articulatory interpretation of these acoustic findings, as described above, is that the apraxic subjects exhibited overscaling of articulatory gestures. If one goal of contrastive stress is to hyper-articulate a given word relative the words around it, then the speakers must be capable of either producing adequate non-local durational effects (i.e. reducing the durations of non-stressed words), or of producing local effects of sufficiently large magnitudes so as to distinguish the stressed word from the surrounding context (i.e. prolonging or
expanding an already overscaled gesture). The perceptual results of these types of modification must be explored; but given the apparent lack of sufficient control of the apraxic subjects in the present investigation to produce these local and non-local changes it would seem that the task of contrastive stress could have a negative influence on the phrase-level rhythm and overall prosody of the utterance. In such cases contrastive stress would be contraindicated.

However, some of the neurogenically impaired speakers did produce, in the contrastive stress condition, the types of temporal and spectral acoustic adjustments that would be expected. Certain dysarthric speakers were able to produce both local and non-local durational effects to distinguish the contrasted word. Furthermore, some speakers were able to produce relatively systematic increases in the F2 transition extent of the stressed word. Whereas the perceptual correlates of these types of acoustic changes remain to be defined, it is expected that these and more general acoustic effects may contribute to the facilitation of listener perception. Such information would support the judicious use of the contrastive stress technique to enhance communication abilities among individuals with motor speech disorders.

Finally, we have suggested several times that the normal and disordered subjects in the present investigation failed to produce consistent patterns of temporal variation as a function of contrastive stress. The lack of consistency, however, may have little to do with variation across disordered subjects, and more to do with the fact that our measures of phrase-level effects of contrastive stress are highly restricted. We provide no data on fundamental frequency (F0) or intensity, two factors that are known to contribute to the production and perception of stress; moreover, there is also the issue of juncture cues, which are similarly ignored here. It is possible that the ‘idiosyncracies’ would disappear if all measures were available, and could be combined into a weighted model of the production of the contrastive stress conditions. For example, perhaps some of the variation in local and non-local effects in apraxic speakers is actually offset by fundamental frequency phenomena, leaving a net result that is more consistent across speakers. The issue of idiosyncracy among disordered and control subjects cannot be completely resolved until the entire set of relevant measures is available, and the appropriate model for combining these factors is determined.

Acknowledgements

This research was supported by a University of Minnesota Graduate School Grant-in-Aid of Research, Artistry and Scholarship and NIH Grant NS18797. We wish to express our appreciation to Kris Tjaden, Stephanie Hanson, and Lisa Heinrich.

References

Contrastive stress 65


WEISMER, G. and MARTIN, R. E. (1992) Acoustic and perceptual approaches to the study of
J. M. Liss and G. Weismer


Appendix: Measurement criteria for durations of vocalic nuclei

Buy Bobby a poppy.

\( \text{al}: \) first glottal pulse following /b/ release to final large-amplitude glottal pulse

\( \text{a}: \) same as /al/

\( \text{a}: \) first glottal pulse following /p/ release to final large-amplitude glottal pulse

preceding /p/ closure

Build a big building.

\( \text{il}: \) first glottal pulse following release of /b/ to final large-amplitude glottal pulse

preceding the flap or /d/ closure

\( \text{il}: \) first glottal pulse following release of /b/ to final large amplitude glottal pulse

preceding /g/ closure

\( \text{il}: \) same as first /il/