Perceptual and Acoustic Analyses of Stress Patterning in Apraxic and Normal Speech

Karen A. Colson, Erich S. Luschei and Linda S. Jordan
The University of Iowa, Iowa City, Iowa

Disturbances in prosody have been described as a primary symptom in apraxia of speech (Rosenbek and Wertz, 1976; Rosenbek, 1980; Kent and Rosenbek, 1982). Perceptually-based descriptions of the prosodic disturbances exhibited by apraxic speakers typically have identified abnormal patterns of word stress as a salient characteristic (Darley, Aronson and Brown, 1975; Rosenbek, 1978; Wertz, LaPointe and Rosenbek, 1984). However, systematic study of the alterations in word stress in apraxic subjects' speech patterns has, as yet, provided only limited descriptive data.

Tonkovich, Asp, Shadden and Mason (1978) analyzed apraxic and normal speakers' ability to reproduce a wide range of prosodic patterns that varied systematically in stress placement. The accuracy of the subjects' responses was judged perceptually without regard to the adequacy of articulation. The apraxic subjects' performance was significantly inferior to that of the normal group, indicating that apraxic speakers exhibit marked difficulty in imitating changes in varied prosodic features, including stress patterning over syllabic sequences.

Presently, all the published quantitative data on altered patterns of word stress in apraxia of speech have come from two studies that broadly examined these prosodic disturbances. In a preliminary study of prosodic disturbance and site of neurologic lesion in apraxia of speech (and three types of dysarthria), Kent and Rosenbek 1982) noted marked alterations in acoustic parameters commonly associated with word stress. Their apraxic subjects' speech patterns were characterized by prolonged intrasyllabic pauses and segments related to phonetic syllable production. The apraxic subjects' syllable durations often were nearly uniform across an utterance and there was lengthening of steady-state segments associated with normally unstressed vowel productions. There also was little intersyllabic dependency in the contour of the apraxic subjects' fundamental frequency variation. According to Kent and Rosenbek, these specific alterations often contributed to a "dissociation" of syllables with no relational pattern across a syllabic sequence. The investigators concluded that perception of abnormal stress in the apraxic subjects' speech could be attributed particularly to a general tendency toward segment prolongation and a failure to effect syllable reduction when appropriate.

In a later study, Kent and Rosenbek (1983) attempted to identify the acoustic correlates of articulatory disturbances in the speech of apraxic subjects. The investigators reported that apraxic subjects' speech was characterized by prolonged duration of syllabic nuclei (steady state segments) associated with normally unstressed vowel sounds and less relative difference, compared with normal speech, between the segment durations related to unstressed and stressed vowels. The apraxic subjects' voice intensity variation was reduced compared with normal subjects and the intensity of unstressed vowels was greater in apraxic speech than in normal speech. Kent and Rosenbek concluded that these acoustically definable alterations in segment duration and intensity variation often resulted in diminished prosodic contrast between syllables in the apraxic subjects' speech patterns.

Currently, the relationship between altered segment duration and the perception of abnormal word stress in apraxic speakers' speech patterns is not well understood. It is unclear if acoustically observable alterations in apraxic speakers' syllable duration patterns typically accompany perceived abnormalities in word stress. Because prosodic disturbance is a predominant feature of apraxia of speech, abnormal stress patterns should be evaluated more thoroughly.

The present study investigated apraxic and normal speakers' speech patterns relative to the perceptual adequacy of articulation and stress placement and the occurrence of associated alterations in syllable duration. Through the use of perceptual and acoustic levels of analysis and stimuli that were controlled relative to changes in segment duration for unstressed and stressed syllables, an answer to the following question was sought.

(1) Do apraxic and normal speakers differ in imitating patterns of reduced and increased segment duration associated with the production of specific unstressed and stressed syllables?

Because the speakers could produce relative changes in fundamental frequency or voice intensity as a function of stress placement, rather than related changes in syllable duration, the acoustic pattern associated with their productions also was examined in order to answer the following question.

(2) What measurable changes in fundamental frequency or voice intensity which may be associated with unstressed versus stressed syllable production are present in apraxic and normal speakers' speech patterns?

METHOD

Individual subject information is presented in Table 1. All subjects were native speakers of English and had normal hearing sensitivity. Normal subjects were matched with aphasic subjects on age and gender. Each apraxic subject had a medically-diagnosed history of left hemisphere brain damage, mild to moderate apraxia of speech (LaPointe and Johns, 1975; Wertz and Rosenbek, 1971; Wertz et al., 1984), no more than mild expressive aphasia, and no significant auditory comprehension deficit (Goodglass and Kaplan, 1972). There was no evidence of structural speech mechanism anomalies, oral paralysis or marked oral apraxia during an oral-peripheral examination and performance of oral nonverbal movement tasks (DeRenzi, Pieczuro and Vignolo, 1966; Darley et al., 1975; LaPointe and Wertz, 1974). There were no reported premorbid communication impairments.

Tape recordings were made as each subject repeated recorded stimuli consisting of the nonsense disyllables, /pəkʌ/, /kətʌ/, and /təpʌ/, embedded in the carrier phrase, "You will say ____." Thirty stimulus phrases were presented once and repeated twice by each apraxic subject and once by each normal subject. A Bruel and Kjaer 2033 High Resolution Analyzer was used to obtain relative measures of duration, fundamental frequency and peak amplitude associated with each phonetic syllable in the subjects' disyllabic productions. Figure 1 illustrates a typical sample of the signal waveform for a disyllabic production. This sample is the waveform for an apraxic subject's (A1) production of /kətʌ/.

Table 1. Subject information. Sex, age, etiology, months post onset, mean rating for oral nonverbal movements, and mean raw scores on Auditory Comprehension Subtests of the Boston Diagnostic Aphasia Examination.

Subject	Sex	Age	Etiology	Time Post CVA	Mean Rating	A.C. Score
Apraxic						
A1	M	60.03	Thrombotic Single CVA	7.5	2.93	28.50
A2	F	48.00	Embolic Single CVA	10.0	3.81	29.75
A3	M	60.07	Thrombo-embolic Single CVA	45.0	3.29	28.25
Normal						
N1	M	60.00				
N2	F	48.02				
N3	M	60.08				

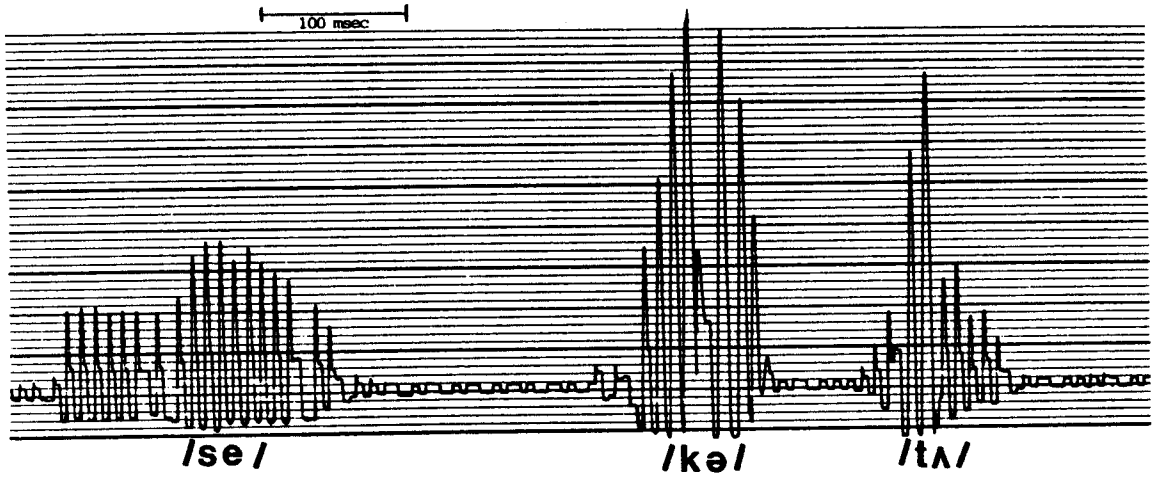

Figure 1. An example of the audio signal waveform for an apraxic subject's production of /kətʌ/.

The disyllabic productions also were analyzed perceptually by three certified speech and language pathologists to determine the accuracy of articulation and stress placement. Based on the perceptual judgments, each

disyllabic production was classified as either accurate or inaccurate for both articulation and stress placement. The acoustic measures for each subject's replications of a disyllable were analyzed statistically according to the accuracy of articulation and stress placement judged to be present in the corresponding productions. Average duration, fundamental frequency, and voice intensity for each syllable were determined for each subject.

Group differences in relative duration, fundamental frequency and voice intensity change for the normal and apraxic subjects' productions judged to be accurate (in both articulation and stress placement) were evaluated with multiple t tests for related measures. Differences in the apraxic subjects' relative duration, fundamental frequency, and voice intensity change in productions judged to have inaccurate stress placement and in utterances perceived by the judges as being accurately stressed (accurate or misarticulated) were evaluated with multiple t tests for related measures.

RESULTS

The results of the statistical tests indicated that there were no significant differences ($p > .05$, $df=5$, directional) between the normal and apraxic groups' average relative syllable duration, fundamental frequency or voice intensity changes for accurate productions. There were, however, some observable trends in the data for the two groups and individual subjects that are noteworthy.

Syllable duration means and standard deviations for each normal and apraxic subject's productions of /pəkʌ/ that were accurate in both articulation and stress placement according to perceptual judgments are presented in Figure 2. Syllable duration means and standard deviations for the model's productions of the disyllable also are presented for comparison. The subjects are identified according to group (N for normal and A for apraxic) and sex (M = male and F = female), and matched normal and apraxic subjects are indicated by number (1, 2, and 3). As can be seen from comparison, the apraxic subjects tended to imitate the model's patterns of duration change associated with word stress less often than the normal subjects. In some instances, individual apraxic subjects evidenced less durational contrast than the normal subjects, as well as demonstrating equalized syllable durations.

Figure 3 depicts the mean duration for each subject's perceptually accurate productions of the unstressed and stressed syllable of /kətʌ/. The data reflect how all of the normal and apraxic subjects' duration patterns correspond to those of the model. It is apparent that each of the normal subject's pattern of duration change is very similar to that of the model. Differences between the normal subjects' and the model's mean duration for the unstressed and stressed syllable range only from 1 to 32 ms and 3 to 33 ms respectively. In comparison, for the apraxic subjects, the duration contrasts produced by A1 are opposite to the model's and A2's and A3's duration changes are similar but tend to be either much smaller (71 ms) or much greater (87 ms) than the model's. Thus, again, overall it is apparent that the apraxic subjects were less consistent than the normal subjects in imitating the duration contrasts.

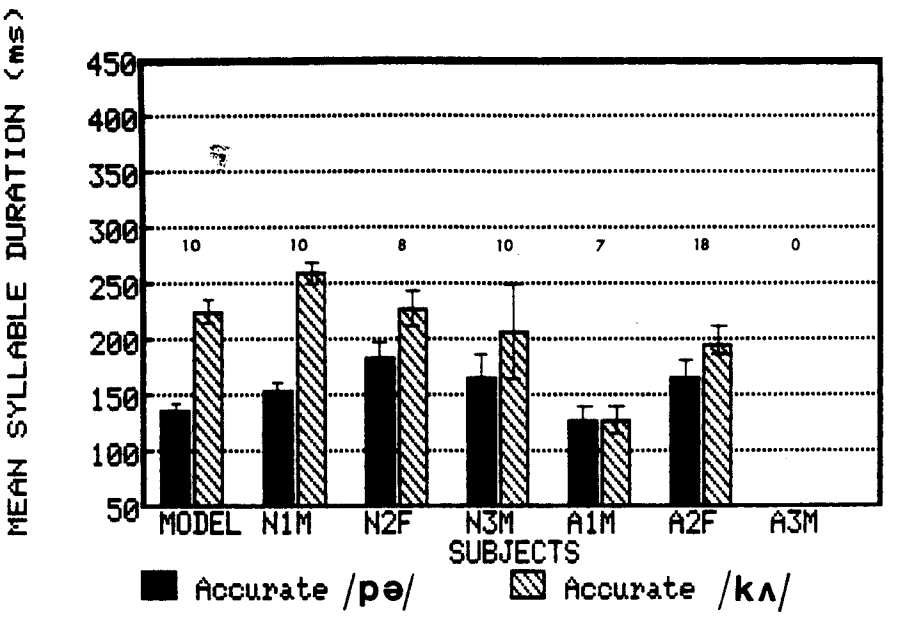

Figure 2. Syllable duration means (in msec) for each normal and apraxic subject's perceptually accurate productions of /pəkʌ/. Means for the unstressed and stressd syllable are indicated by the first and second bar. Line bars indicate plus and minus one standard deviation. The number of productions each set of means is based on is shown above the data bars. Subjects are identified by group, matching and sex.

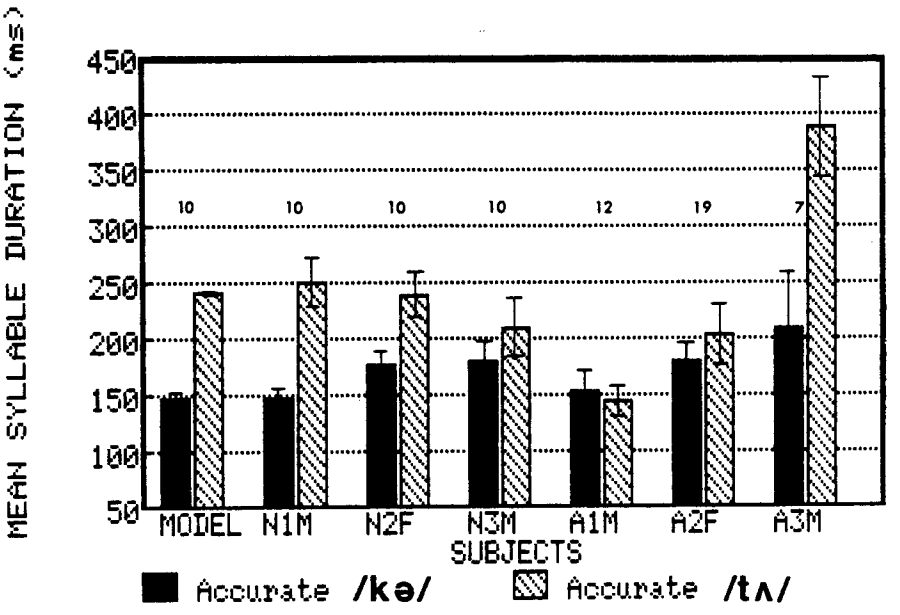

Figure 3. Syllable duration means (in msec) for each normal and apraxic subject's perceptually accurate productions of /kətʌ/. Refer to caption of Figure 2 for a description of symbols.

All subjects tended to change fundamental frequency in relation to syllable stress, but the apraxic subjects often produced a greater contrast than the normal speakers. An example of the pattern of change in fundamental frequency produced by each subject is illustrated in graphic form in Figure 4 The upper portion of the figure shows the fundamental frequency means for each normal and apraxic subject's perceptually accurate productions of /k t / The ratio means in the lower portion of the figure were determined from the same data and reflect each subjects' average fundamental frequency for the unstressed syllable relative to that for the stressed syllable. Overall, the average amount of relative change in fundamental frequency produced by each subject often appeared to vary inversely from the average magnitude of duration change produced in the same utterances. Note, for example, that the greatest relative change in fundamental frequency in relation to stressed syllable production was by apraxic subject A1, who tended to demonstrate a pattern of equalized duration in producing the disyllables.

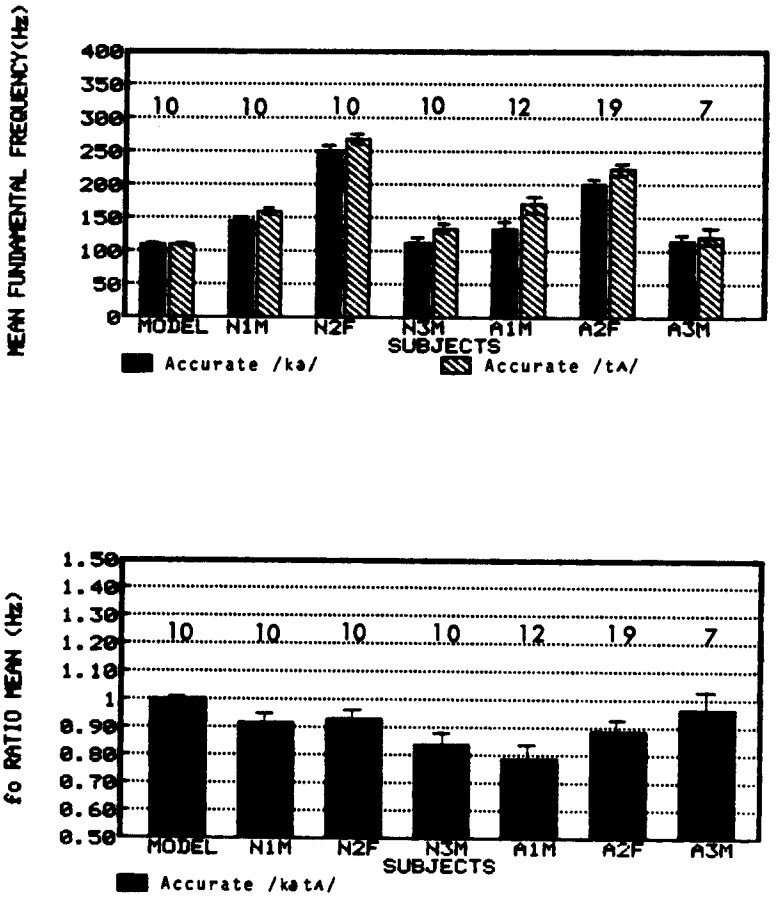

Figure 4. Mean fundamental frequency (upper portion) and fundamental frequency (fo) ratio mean for each subject's perceptually accurate productions of the unstressed and stressed syllable in /kətʌ/. Line bars indicate plus and minus one standard deviation. The number of productions each mean is based on is shown above the data bar.

Both the normal and apraxic groups tended to produce unstressed syllables with greater voice intensity relative to stressed syllables. Thus, a large increase in voice intensity was not one of the ways any of the subjects stressed syllables in their accurate disyllabic productions.

The apraxic subjects' relative duration and voice intensity changes in utterances judged to have inaccurate stress placement did not differ significantly ($p > .05$, $df=9$, directional) from the changes produced in utterances with accurate stress placement (accurate or misarticulated). There was, however, significantly less ($p < .05$, $df = 9$, directional) fundamental frequency change (from first to the second syllable) in apraxic utterances judged to be inaccurately stressed than in those perceived by the listeners to have accurate stress placement. These trends often were less apparent in the data for individual apraxic subjects. Typically, a combination of reduced contrasts in the acoustic parameters was associated with each apraxic subject's productions that were judged perceptually to have inaccurate stress placement.

DISCUSSION

In the present study, relative differences between the durations for unstressed and stressed syllables tended to be smaller in the apraxic subjects' speech patterns than in the normal subjects. There also were instances in which one or more of the apraxic subjects tended to equalize the duration of unstressed and stressed syllables. Similar findings were reported previously by Kent and Rosenbek 1982, 1983). However, Kent and Rosenbek concluded that these duration patterns contributed in part to the perception of abnormal stress in their apraxic subjects' speech. The results of the present study partially support this conclusion but also suggest that these duration patterns are not always associated with the perception of abnormal stress in apraxic subjects' speech. The data indicate that these duration patterns were present in the acoustic signal for apraxic subjects' productions that were judged perceptually to be accurate as well as those judged inaccurate with regard to stress placement. Often there was less duration contrast between the acoustic segments defined for the syllables in the inaccurately stressed productions, but the same type of duration patterns were associated with utterances that were judged to differ with regard to the accuracy of stress placement. Thus, it may be inferred that reduced duration contrasts (for unstressed vs stressed syllables), including equal syllable duration, in apraxic speakers' speech patterns may not necessarily be related to the perception of inaccurate stress placement in their utterances.

In summary, it may be concluded from the results of this study that apraxic and normal groups of speakers do not differ significantly in imitating patterned changes in syllable duration, fundamental frequency, or voice intensity associated with the production of unstressed and stressed nonsense syllables. It also may be concluded that there are significant differences in relative fundamental frequency changes associated with apraxic speakers' imitative productions of nonsense disyllables perceived by listeners to be either accurate or inaccurate with regard to a particular pattern of stress placement. Future study of apraxic speakers' production and perception of varied pitch and timing changes associated with different stress patterns may provide increased insight as to the nature, evaluation and treatment of abnormal stress patterning in apraxia of speech.

REFERENCES

Darley, F.L., Aronson, A.E., and Brown, J.R. Motor Speech Disorders. Philadelphia: W.B. Saunders Company, 1975.

DeRenzi, E., Pieczuro, A., and Vignolo, L. Oral apraxia and aphasia. Cortex, 2, 50-73, 1966.

Freeman, F.J. Prosody in perception, production and pathologies. In N.J. Lass, L.V. McReynolds, and D.E. Yoder (Eds.), Speech, Language and Hearing: Pathologies of Speech and Language. Philadelphia: W.B. Saunders Company, 1982.

Goodglass, H. and Kaplan, E. Boston Diagnostic Aphasia Examination. Philadelphia: Lea and Febiger, 1972.

Huggins, A.W.F. Just noticeable differences for segment duration in natural speech. Journal of the Acoustical Society of America, 51, 1270-1278, 1972.

Kent, R.D. and Rosenbek, J.D. Acoustic patterns of apraxia of speech. Journal of Speech and Hearing Research, 26, 231-249, 1983.

Kent, R.D. and Rosenbek, J.D. Prosodic disturbance and neurologic lesion. Brain and Language, 15, 259-291, 1982.

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

LaPointe, L.L. and Johns, D.F. Some phonemic characteristics in apraxia of speech. Journal of Communication Disorders, 8, 259-269, 1975.

LaPointe, L.L. and Wertz, R.T. Oral-movement abilities and articulatory characteristics of brain-injured adults. Perceptual and Motor Skills, 39, 39-46, 1974.

Rosenbek, J.D. Apraxia of speech -- relationship to stuttering. Journal of Fluency Disorders, 5, 233-253, 1980.

Rosenbek, J.C. Treating apraxia of speech. In D.F. Johns (Ed.), Clinical Management of Neurogenic Disorders. Boston: Little, Brown and Company, 1978.

Rosenbek, J.C. and Wertz, R.T. Veterans Administration Workshop on Motor Speech Disorders. Unpublished, Madison, Wisconsin, 1976.

Tonkovich, J.D., Asp, C.W., Shadden, B.B. and Mason, D. Imitation of supra-segmental patterns by apraxic and normal adult speakers. Unpublished Paper presented to the Tennessee Speech and Hearing Association Convention, Chattanooga, Tennessee, 1978.

Wertz, R.T., LaPointe, L.L. and Rosenbek, J.C. Apraxia of Speech in Adults The Disorder and its Management. Orlando: Grune and Stratton, Inc., 1984.

Wertz, R.T. and Rosenbek, J.D. Appraising apraxia of speech. Journal of Colorado Speech and Hearing Association, 5, 18-36, 1971.

DISCUSSION

Q: I wonder if you could give me a little clearer idea of just where we think we're going with these kinds of studies. I just am not sure of what direction these studies are taking us or how it's going to help me when I go back to my clinic.

A: I think it is important for us to study apraxia of speech on a physiologic level, on an acoustic level, and to continue with perceptual studies. With particular regard to acoustic studies, I think it

is relevant to realize that we might be able to identify abnormalities when we compare apraxic and normal speakers' acoustic patterns, in that apraxic speakers' patterns may be altered or different from those of normal speakers, but when we listen to their speech, perceptually we may not hear those differences. Ultimately, we are working toward developing functional communication for these patients. While there may be identifiable abnormalities in the acoustic pattern of apraxic speakers' speech, it is questionable how relevant these are to what we hear and do in therapy. Just because apraxic speakers' acoustic patterns differ from normal speakers' doesn't necessarily mean that the patterns are inadequate or warrant treatment. I think it is relevant to continue to pursue analysis of apraxic speakers' speech behavior and patterns of behavior on all levels, because all levels have something to contribute. We can infer about motor programming, where the breakdowns may be, and how this all may relate to a linguistic versus a motoric basis for errors. Personally, I'm very interested in the basis of the prosodic disturbances in apraxia of speech, and I don't necessarily believe they are compensatory in nature. I think it is only through different levels of analysis that we're going to be able to clearly define the basis of the disturbances and make appropriate inferences in relation to motor programming.

C: Even though we have the focus of application in this conference and everyone is always trying to relate things to treatment, there are a variety of reasons for doing research and scientific inquiry. One of them is the level of explanation, as well as the level of application. Personally, I could use all the explanation I can get, particularly in this disorder, and I think you pointed that out.

C: I have a comment about levels of explanation just mentioned. I don't think apriori that acoustic or physiologic, or any other level including nerve conduction velocity, is an explanation or any closer to an explanation. So, I suggest caution the use of the term "explanation" without relating it to an explanation of a theory or a model.

C: I did not mean to imply that levels of explanation were not without interpretation. Pages of data by themselves are islands.

Q: When your judges listened to the productions and judged the accuracy of them in terms of stress and the phonetic content, did you see any differences in those judgments in that the subjects made fewer or more errors of a certain type?

A: Based on the judgments made in the perceptual analysis, a consistent error pattern was not evident in the responses produced by the group of apraxic subjects. All of the apraxic subjects exhibited errors of stress placement in producing the three disyllables. However, the number of productions judged to be inaccurately stressed appeared to vary greatly across subjects and the stimulus items. The productions of apraxic subject A2 were judged to be inaccurate only with regard to stress placement, while the errors judged to be present in the productions of A1 and A3, who evidenced a more severe output impairment, varied considerably across the three error categories. Many of their errors involved both misarticulation and inaccurate stress placement.

Q: Did you plan to make the nonsense syllables a little more nonsensical in that you used a stress pattern that is not typical in English in the model, as you stressed the last syllable instead of the first one? If you did do that, did you do that on purpose?

A: The stimuli that I used were designed particularly for the task with regard to acoustic segmentation and phonetic context. I tried to select consonant sounds that were less likely to evoke articulation errors. The stress pattern that I used is not typical in English, except for contrasting the production of some adjective/noun and noun/verb pairs. I did not select the pattern for a special reason. Although, in addition to the stimuli being nonsense, since the stress pattern was not typical, it did provide a type of stimulus that perhaps allowed me more control in examining if the apraxic subjects changed duration in relation to stress production and how they differed from normal speakers in doing so.