SPEECH MOTOR CONTROL IN ATAXIC DYSARTHRIA

P.F. McCormack*, J. C. Ingram~

*Department of Speech Pathology
Flinders University
South Australia

~Department of English
University of Queensland

ABSTRACT - Ataxic dysarthria has unique prosodic characteristics in which the production of linguistic stress and speech rhythm are disturbed. Words and syllables are usually produced at an extremely slow rate, with long pauses between them. There is often the perception that words and syllables are being produced with "equal stress". Kent, Netsell & Abbs (1979) suggested that this disturbed prosodic expression may reflect an impairment in anticipatory motor programming as well as any associated difficulties with motor execution. An experiment is reported on the production of "stress shifts" requiring anticipatory planning by 10 speakers with ataxic dysarthria, compared to 10 matched speakers with normal speech production. Comparison with normal speakers at normal and slow rates of speech indicates that the speakers with ataxic dysarthria, unlike the control subjects, do not take account of the position of the main stress in the following word in their production of the "shift" words. This pattern is consistent with a disruption to motor programming, and is consistent with current knowledge about the functions of the cerebellum in motor control.

INTRODUCTION

Ataxic dysarthria is the name given to the speech characteristics that result from damage to the cerebellum or the cerebellar pathways that link the cerebellum to the cortex, to the basal ganglia,and to the spinal cord (Darley, Aronson & Brown, 1975). The changes in speech that occur with cerebellar damage are particularly distinctive compared to those found in other types of dysarthria. Changes in the expression of linguistic stress and speech rhythm have been the most salient feature consistently reported since the first studies in the late 19th century. Most studies describe a slow and regular speech rhythm, with the careful production of each word, and sometimes each syllable. Often there are noticeable pauses between words and syllables. This gives the impression to the listener of the person speaking with an "even" stress such that each word and syllable was given equal prominence. This unique prosodic pattern has often been labelled as "scanning speech".

The role of the cerebellum in motor control is not fully understood, but it appears to be involved in the production of rapid, fluent and coordinated movements. Two principal functions have been identified: the "moment-to-moment" regulation of planned movements as they are being executed, and the motor programming of sequences of planned movements in concert with cortical areas (Eccles, 1979; Gracco & Abbs,1987). Llinas (1985) proposed that the lateral zones of the cerebellum convert coarsely specified pre-motor cortex specifications for movements into specific spatial and temporal coordinates to be achieved in motor execution. The cerebellum is thought to do this by providing a predictive "map" of the timing and spatial requirements for the particular movement sequence. Gracco & Abbs (1987) describe this process of revising motor commands as one "...whereby on-line sensory input and general motor command prespecifications are "mixed" dynamically to yield appropriate intended goals." (p. 175). Impairment to the cerebellum's modifying function in motor programming would result in the imprecise motor commands from the motor cortex not being adjusted, or at least effectively adjusted, to take account of the overall movement sequence to be achieved and the current peripheral sensory information to guide the accurate achievement of those intended articulatory goals. While movements would still occur they would lack the fine-grained attunement to the particular timing and spatial targets required for the overall movement sequence to be fluent.

Kent, Netsell, & Abbs (1979), and Ingram, Murdoch & Chenery (1988) have suggested that the presenting pattern of prosodic disturbance in ataxic dysarthria could arise from either difficulties in the cerebellar regulation of the ongoing execution of a movement sequences, or from an impairment to the cerebellum's role in the anticipatory motor programming of speech movement sequences. Kent et al. (1979) speculated that this "scanning speech", where words and syllables are slowed down, and are segregated by long pauses, could reflect a voluntary or involuntary strategy by the ataxic speaker to maximise the time available for peripheral feedback to assist in regulating the movement of the articulators. Alternatively, the same pattern could reflect a voluntary or involuntary strategy to limit the amount of input at any one time from the pre-motor phonological plan to the "articulatory buffer" that assembles the motor program. In severe cases, each word or syllable would be programmed and produced without reference to the following word in the utterance. In order to distinguish between these 2 quite distinct disruptions to speech motor control, what is required is the investigation of an aspect of prosody that is dependent on sequential lookahead for its functioning.

One aspect of English prosody dependent on sequential lookahead is the rhythm rule. The stress pattern of some English words in connected speech is dependent on the stress pattern of the following word. It is generally acknowledged that these shifts in the prominence pattern on some words are due to a strong rhythmic constraint to prefer the alternation of stressed and unstressed syllables and words, and to avoid the juxtaposition of stresses (Liberman & Prince, 1977). Hence, while *ornate* spoken in a noun phrase such as *the ornate one* has the main stress on the last syllable, in a noun phrase such as *the ornate cup* there is a perception that the main stress has shifted to the first syllable of *ornate*. This rhythm rule can be formulated as an operation where stress shifts from one syllable of a word on to another in order to avoid "clashing" with an adjoining stress. Selkirk (1984) provides a formulation to display these stress shifts on a "metrical grid". Each beat of prominence in the utterance can be represented by an asterisk (*). These prominence beats form a layered hierarchy (grid), with each layer representing another level of prominence within the utterance. On the first layer each syllable is represented by a beat, on the second layer, each stressed syllable, on the third each word's main stress. Beats can shift leftward within the metrical grid so as to avoid a "stress clash" of 2 beats adjoining each other. A metrical grid representation of the 2 *ornate* phrases could be as follows:

```
              *                                *
              *                     *          *
        *     *                     *     *    *
  *     *     *     *         *     *     *    *
 the    or   nate  one       the    or  nate  cup
```

AIM

The aim of this experiment is to determine whether a disturbance in the anticipatory speech motor programming of linguistic stress shifts can be implicated in speakers with ataxic dysarthria. The production of the rhythm rule will be compared between subjects with ataxic dysarthria and matched controls with normal speech production. It is predicted that if anticipatory motor programming has been affected, then the speakers with ataxic dyarthria will not take account of the stress pattern in the word following any potential shift word. On the other hand, general difficulties with the timing and spatial regulation of articulatory movements should result in intact, but less precise productions of the rhythm rule than the control subjects.

METHOD

Ten subjects with multiple sclerosis, and who were native speakers of English, were selected. They had been diagnosed by a neurologist as exhibiting predominantly ataxic symptoms, and by an experienced speech pathologist as exhibiting ataxic dysarthria with "scanning speech" prosody. No subject had a prior history of any other type of speech or language difficulty. These were matched for age and gender with 10 native speakers of English who had normal speech production and no history of speech and language problems. All subjects were assessed as having good hearing, and reading comprehension ages equivalent to an upper high school level. Each subject was recorded reading a series of sentences containing noun phrases which comprised of a potential stress shift word followed by a word with varying syllable distance to its main stress. The sentences were designed to provide a phonological context where shift and non shift environments could be manipulated. Examples of the 5 contexts used are as follows:

Two contexts where no shift was predicted:

No stress following	There were thirteen of them.
Shift word focused	There were THIRTEEN officers at the party.

Three contexts where shift was predicted:

One syllable distance	There were thirteen 'officers at the party.
Two syllable distance	There were thirteen o'fficials at the party.
Three syllable distance	There were thirteen poli'ticians at the party.

Six potential stress shift words were used: *thirteen, bamboo, sardine, underdone, overnight*, and *japanese*. These words had been identified in a previous experiment as being particularly susceptible to stress shift in speech production. Each word consisted of 2 metrical feet, 3 had feet containing one syllable each (eg I bam I boo I), while 3 had feet with the first foot containing 2 syllables and the second one syllable (eg I japa I nese I).

Analysis

Recorded shift words, embedded in their noun phrases, were digitised at 20.8 kHz using the Soundscope speech signal processing program. The duration of the shift word, the duration of each foot, and the duration of the pause between the shift word and the following word was measured. In order to obtain a measure of variation in the duration of the first foot compared to the second foot, the duration of the first foot as a percentage of the duration of the whole word was calculated (henceforth described as relational duration). The peak fundamental frequency for each foot was also calculated using a peak-picking algorithm within the Soundscope program. In order to obtain some measure of the relative changes in fundamental frequency pattern between the 2 feet over different contexts, the value for the second peak was subtracted from that of the first (henceforth described as fundamental frequency shift). Because of the expected absolute durational differences between the ataxic speakers and their controls, the duration of any pause between the shift and fulcrum words was expressed as a percentage of the duration of the whole noun phrase. The rate of speech for each subject was calculated by dividing the duration of each sentence by the number of syllables it contained. The recordings for the ataxic and control groups were combined and edited on to an audio tape in a pseudo-randomised order. Three phonetically trained linguists were asked to rate the stress levels in each shift word token as either: 1) the last stressed syllable is more prominent 2) both stressed syllables have equal prominence, or 3) the first stressed syllable is more prominent. These perceptual judgements were converted to a numeric expression of stress shift: zero for no shift, one for equal prominence, and 2 for full shift.

RESULTS

Both the ataxic group and the control group were perceived as shifting stress in the 3 Rhythm contexts, and not shifting in the 2 non rhythm contexts, though the range of perceived shift variation was narrower for the ataxic group over the 5 contexts. The 3 judges reported that they found making stress placement judgements for the ataxic group more difficult than for the control group because of the general prosodic distortion. A 2 way analysis of variance for group membership and context against the judges' perception of stress shift indicated that there were significant main effects for both group and context as well as a significant 2 way interaction between them (context: $p = .000$, $F = 427.3$, d.f = 4, 596; group: $p = .000$, $F = 50.7$, d.f. = 1, 599; interaction: $p = .000$, $F = 24.5$, d.f. 4, 596). The most important difference between the 2 groups was in the pattern of stress shift over the 5 contexts. While the control group displayed a graded decrease in the perception of shift as the syllable distance to the main stress in the fulcrum word increased, the ataxic group showed no such systematic pattern. The ataxic group were perceived as shifting to the same extent in all 3 Rhythm contexts. Figure 1. displays an error bar plot with a 95% confidence interval of the average stress shift judgement for each of the 5 contexts for the ataxic and control groups.

There were acoustic-phonetic changes in the shift words that corresponded to these perceived patterns of stress shift in both groups. A 2 way analysis of variance for group membership and context against the relative duration of the first foot in each shift word indicated that there was a significant main effects for context but not for group membership ($p = .000$, $F = 67.1$, d.f. 4, 596). As well there was a significant 2 way interaction between context and group ($p < .05$, $F = 3.4$, d.f. = 4, 596). Both the ataxic and control groups had similar significant increases of approximately 9% in the relative duration of the first foot in the Rhythm contexts compared to the 2 non rhythm contexts. However, the pattern of increase in relative duration was not the same for the 2 groups. The control group displayed a graded decrease in the relative duration of the first foot as the syllable distance to the main stress in the fulcrum word increased. The ataxic group showed no such systematic difference across the 3 Rhythm contexts, with the degree of durational change remaining the same. Figure 2 displays an error bar plot with a 95% confidence interval of the mean relative duration of the first foot across the 5 contexts for the control and ataxic groups. There were similar results for the shift in peak fundamental frequency between the 2 feet in each word. A 2 way analysis of variance for context and group membership against shift in peak fundamental frequency between the 2 feet indicated significant main effects for both context and group, as well as a significant interaction between them (context: $p = .000$, $F = 57.6$, d.f. = 4, 596; group: $p < .01$, $F = 7.4$, d.f. = 1, 599; interaction: $p = .000$, $F = 11.5$, d.f. = 4, 596). For both the control and ataxic groups there were significant positive shifts in peak fundamental frequency in the Rhythm contexts compared to the non rhythm contexts. The ataxic group differed from the control group, however, in having no graded decrease in peak fundamental frequency shift as the syllable distance increased. Figure 3 displays an error bar plot with a 95% confidence interval of the mean shift in peak fundamental frequency between the 2 feet across the 5 contexts for the control and ataxic groups.

Because the ataxic group spoke at such a slow rate of speech, a further comparison was made with the production of the rhythm rule by speakers with normal speech production speaking at a slow tempo. (For a fuller description of the production of the Rhythm Rule at a slow tempo refer to the paper *Tempo and the Rhythm Rule* in these proceedings). It may be that the difference in pattern between the ataxic and control groups was a reflection of their differences in rate of speech rather than any impairment to motor control as such. The was a marked difference in the rate of speech between the control and ataxic groups. The rate of speech for the control group was 4.8 syllables per second (s.d. 1.4), while for the ataxic group it was 2.0 syllables per second (s.d. 0.9). The pattern of pause duration to word duration, however, was not the same across the 2 groups. The average total duration of the shift words for the control group was 0.5 seconds (s.d. 0.1) with that for the ataxic group being 1.1 seconds (0.4 s.d.). This maintains an approximate 2:1 word duration ratio between the 2 groups. However, the average duration of the pause between the shift and fulcrum words for the control group was 0.02 seconds (0.03 s.d.) with that for the ataxic group being 0.33 seconds (0.38 s.d.). This reflects an approximate 17:1 pause duration ratio between the 2 groups. A comparison with the word and pause duration of normal speakers undertaking the same task at a slow tempo indicated that the ataxic group's results do not simply reflect the characteristics of speaking at a slow rate. The normal speakers at the slow tempo had a syllable rate of 2.4 syllables per second, which is very similar to that for the ataxic group. Their average total duration of the shift words was also comparable to the ataxic group at 0.91 seconds. However, the duration of the pause between the shift and fulcrum words at the slow tempo for the normal speakers was only 0.03 seconds (0.03 s.d.). This is an approximate 11:1 difference in pause duration between the ataxic group and normal speakers at an equivalent rate of speech. The perceptual and acoustic-phonetic characteristics of the rhythm rule produced by normal speakers at a slow tempo are markedly different from the ataxic productions. The ataxic group were perceived as shifting in all Rhythm contexts, and to the same extent across those contexts. The normal subjects at the slow tempo were not perceived as shifting, though they still preserved some differential effect in response to increasing the number of syllables to the main stress in the following word.

In summary, the ataxic group did not perform in the same way in this experiment as speakers with normal speech production. At a slow rate of speech, speakers with normal speech production do not undergo shifts in stress, while at a normal rate they do. The ataxic group, although speaking at a slow rate, did undergo stress shifts in their speech. However, the perceptual an acoustic-phonetic pattern of stress shifting for the ataxic group was different from that for the control group. Unlike the control group, the ataxic group did not exhibit either perceptual or acoustic-phonetic signs of a decreasing gradation in stress shift as the number of syllables to the main stress in the following word increased.

DISCUSSION

The results for the ataxic group indicate that they underwent stress shift in the appropriate contexts, but were unable to make graded phonetic adjustments to the number of syllables to the main stress in the following word. While there was evidence that the ataxic group could anticipate that the following word was stressed, they were not able to make allowances in the production of the shift word for the phonological characteristics of the following word. This result could not be accounted for by their slow rate of speech, since speakers with normal speech production do not undergo stress shift at all at an equivalently slow rate of speech. Nor could it be accounted for by a disruption to sensory feedback to the cerebellum about the present status and position of the articulators. Stress shifting relies on anticipated knowledge of the movement sequence yet to occur. The results are consistent with an impairment to the cerebellum's role in the anticipatory motor programming of an utterance, which is to adjust the imprecise motor commands from the motor cortex to take account of current and upcoming circumstances so as to accurately achieve intended articulatory goals. While stress shifts still occur, they lack the fine-grained attunement to the particular timing and fundamental frequency targets required in relation to the motor sequence for the upcoming word. The results of this experiment suggest that the traditional classification of the speech disorder that arises from cerebellar or cerebellar tract damage as a dysarthria (and therefore not a speech programming disorder) is highly questionable.

REFERENCES

Darley, F., Aronson, A., & Brown, J. (1975). *Motor Speech Disorders. Philadelphia*: Saunders.

Eccles, J.C. (1979). Introductory remarks. In J. Massion & K. Sasaki (Eds.). *Cerebro-Cerebellar Interactions*. pp. 1-18. Amsterdam: Elsevier / North-Holland Biomedical Press.

Gracco, V.L. & Abbs, J.H. (1987). Programming and execution processes of speech movement control: Potential neural correlates. in E. Keller & M. Gopnik (Eds.) *Motor and Sensory Processes of Language*. pp. 163-201. New Jersey: Lawrence Erlbaum Associates.

Ingram, J., Murdoch, B., & Chenery, H. (1988). Prosody in hypokinetic and ataxic dysarthria. *Proceedings of the First Australian Conference on Speech Science and Technology*. (pp. 350-355). Canberra: Australian National University.

Kent, R.D., Netsell, R., & Abbs, J.H. (1979). Acoustic characteristics of dysarthria associated with cerebellar disease. *Journal of Speech and Hearing Research, 22*, 627-648.

Liberman, M. & Prince, A.(1977) *On stress and linguistic rhythm*, Linguistic Inquiry, 8, 249-336.

Llinas, R. R. (1985). Functional significance of the basic cerebellar circuit in motor coordination. In J.R. Bloedel, J. Dichgans, & W. Precht (Eds.). *Cerebellar Functions*. (pp. 170-185). Berlin:Springer-Verlag.

Selkirk, E. (1984) *Phonology and syntax: The relationship between sound and structure*, (M.I.T. Press: Cambridge, Mass.).

Figure 1. An error bar plot with 95% confidence intervals of mean stress shift judgement across the 5 contexts for the control and ataxic groups.

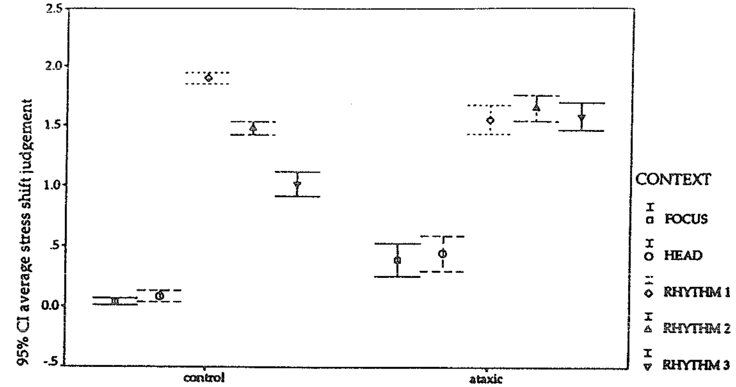

Figure 2 An error bar plot with 95% confidence intervals of the mean relative duration of the first foot across the 5 contexts for the control and ataxic groups (in seconds).

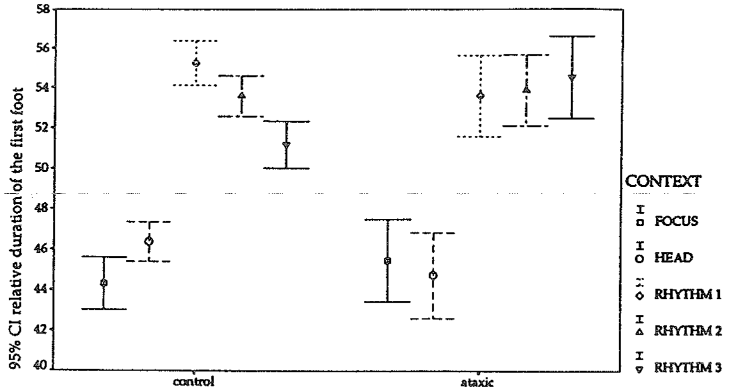

Figure 3. An error bar plot with 95% confidence intervals of the mean shift in peak fundamental frequency between the 2 feet across the 5 contexts for the control and ataxic groups (in Hertz).