THE RHYTHM RULE IN SPEECH PRODUCTION: THE EFFECT OF INTER-STRESS SYLLABLES

P.F. McCormack, J. C. Ingram~*
**Department of Speech Pathology, Flinders University, South Australia, &*
~Department of English, University of Queensland

ABSTRACT

The adjustment of linguistic stress patterns under the influence of rhythm is well attested, though the effects on speech production have been little investigated. An experiment is reported on the perceptual and acoustic effects on the production of the Rhythm Rule of manipulating the number of syllables between primary stresses.The tendency for stress shift to occur decreased as the number of syllables between primary stresses increased. There were both fundamental frequency and durational changes involved in the perceived shifts.

INTRODUCTION

It is generally acknowledged that shifts in the prominence patterns on some words in connected speech are due to a strong rhythmic constraint to prefer the alternation of stressed and unstressed elements [1] [2]. While *bamboo* spoken in a noun phrase such as *the bamboo one* has the main stress on the last syllable, in *the bamboo chair* there is a perception that the main stress has shifted to the first syllable. The Rhythm Rule formulated by Selkirk [2] involved a formal operation where stress shifts from one syllable of a word on to another in order to avoid "clashing" with an adjoining stress. Gussenhoven [3] proposed an alternative formulation whereby the perceived change in prominence is due to a process of pitch accent deletion rather than of phonetic stress "shift", at least in the pre-nuclear position. Horne [4] investigated Gussenhoven's [3] hypothesis in a single speaker of English, and found that the primary phonetic correlate for the Rhythm Rule was a decrease in the fundamental frequency in the second stressable syllable. These results were consistent with an accent deletion formulation. The primary phonetic cue for "stress shift" was change in fundamental frequency on the second stressable syllable of the "shift" word.

However, Horne [4] only investigated phrases in which the potential shift word was followed by a word with its primary stress on the first syllable (for example, *Dundee tartan*). What has not been investigated as yet is whether the tendency for the Rhythm Rule to occur is also dependent on the number of syllables between the primary stresses in both the "shift" word and the following word. For example, is "stress shift" more likely in *bamboo chair* than it is in *bamboo decoration*? If syllable number does play a role in the rhythm rule, it may also follow that durational changes in the "shift" words also provide a phonetic cue for "stress shift".

AIM

The aim of this experiment is to investigate the perceptual and acoustic effects on the production of the Rhythm Rule of altering the number of syllables between the main stress in the potential stress shift word and that in the following word (fulcrum).

PROCEDURE

Fifteen speakers of Australian English were recorded reading a series of sentences containing noun phrases which comprised of a potential stress shift word followed by words with varying syllable number to their primary 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: *They were japanese ones at the hotel.*

Shift word focused: *They were JAPANESE tourists at the hotel.*

Three contexts where shift was predicted:

One syllable: *They were japanese tourists at the hotel.*

Two syllables: *They were japanese developers at the hotel.*

Three syllables: *They were japanese politicians at the hotel.*

Six potential stress shift words were used: *thirteen, bamboo, sardine, underdone, overnight,* and *japanese.* These words had been identified in a previous experiment [5] as being particularly susceptible to stress shift in speech production.

ANALYSIS

Recorded shift words, embedded in their noun phrase, 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 (relative 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 (fundamental frequency shift).

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.

RESULTS

The perceptual results indicated that not only was there a strong perception of shift in the 3 rhythm contexts, but the strength of the shift dropped away as the syllable number between the main stress in the shift word and the main stress in the following word increased. There was a clear pattern to the perceived stress shift judgements across the contexts. The 3 contexts in which shift is predicted demonstrate strong shift values, well above the "equal prominence" value of one. The 2 contexts in which shift is predicted not to occur demonstrate low shift values, well below the value of one. A one way analysis of variance for context against stress shift judgement with an adjusted least significant difference (Bonferroni) test at the .05 level indicated significant differences between the rhythm and non rhythm contexts, and between the Rhythm 3 context (with 3 syllables distance) and the other 2 rhythm contexts ($p = .000$, $F = 178.28$, d.f. = 4, 375). Figure 1 displays the mean perceptual stress shift ratings for each context.

In all subjects there were phonetic changes in the shift words that corresponded to the judgements of stress shift. In the rhythm contexts the relative duration of the first foot was higher than in the non-rhythm contexts. There was also a positive increase in

fundamental frequency shift. Inspection of the data indicated that for 11 of the 15 subjects these results were due not to absolute changes in the first foot of each word but to changes in the second. The absolute duration and fundamental frequency of the second foot decreased in the rhythm contexts. This resulted in the relative prominence between the 2 feet shifting from the second to the first. However, for 4 of the subjects there were changes in the absolute duration and fundamental frequency for the first foot. Figure 2 displays the mean values for percentage durational change for each context, while Figure 3 displays the mean fundamental frequency shift for each context.

One way analyses of variance for relative duration of the first foot against context, and for fundamental frequency shift against context with adjusted least significant difference (Bonferroni) tests at the .05 level indicated significant differences between the rhythm and non rhythm contexts, and between the Rhythm 3 context (with 3 syllables distance) and the other 2 rhythm contexts ($p = .000$, $F = 29.1$, d.f. = 4, 375; $p = .000$, $F = 34.8$, d.f. = 4, 375). A step-wise linear regression analysis indicated that changes in relative duration was the primary acoustic correlate for the judges' perception of stress shift, followed by shifts in peak fundamental frequency.

DISCUSSION

The results provide confirmation that the Rhythm Rule is dependent on the metrical structure of the word following the shift word. In particular, the number of syllables between the primary stresses of the 2 words is critical for the expression of the Rhythm Rule. The suggestion by Gussenhoven [3] and Horne [5] that the phonetic realisation of the Rhythm Rule is one of pitch accent deletion needs qualification. While it certainly involves a decrease in fundamental frequency on the second foot in the shift word for most speakers, there were also systematic durational adjustments dependent on the metrical structure. The regression analysis indicated that these durational adjustments were the primary phonetic cue used by listeners.. Horne's [4] results, suggesting the primacy of fundamental frequency as the cue for "shift" most likely reflect the effects of investigating the Rhythm Rule without taking variations in inter-stress syllable number into account.

The increased numbers of subjects in this study also highlighted that for some subjects there were positive phonetic changes on the first foot in rhythm contexts. For these subjects, a deletion model of the Rhythm Rule does not appear to provide an appropriate description.

REFERENCES

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

[2] Selkirk, E. (1984) *Phonology and syntax: The relationship between sound and structure*. Camb., Mass.: MIT Press.

[3] Gussenhoven, C. (1986) Review of Selkirk 1984. *Journal of Linguistics*, 22, 455-474.

[4] Horne, M. (1990) Empirical evidence for a deletion formulation of the rhythm rule in English. *Linguistics*, 28, 959-981.

[5] McCormack, P. & Ingram, J. (1989) Phonetic evidence for the rhythm rule. Paper presented at *The Australian Language & Speech Conference*, Melbourne: Monash University.

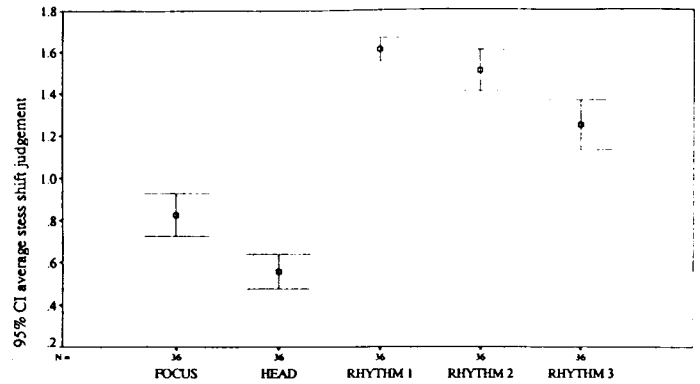

Figure 1. An error plot with 95% confidence interval of the mean stress shift judgement for each of the 5 contexts.

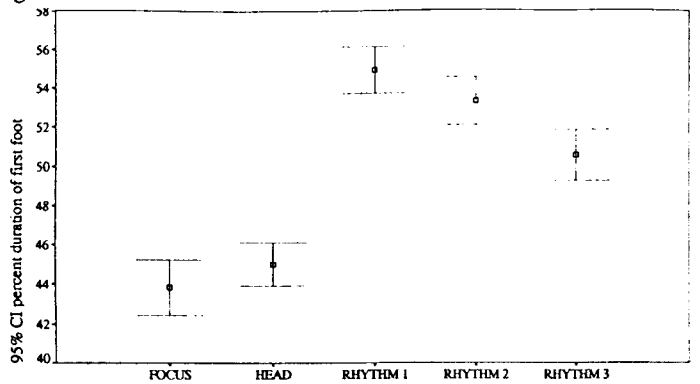

Figure 2. An error plot with 95% confidence interval of the mean relative duration of the first foot in the shift words for each of the 5 contexts (as a percentage).

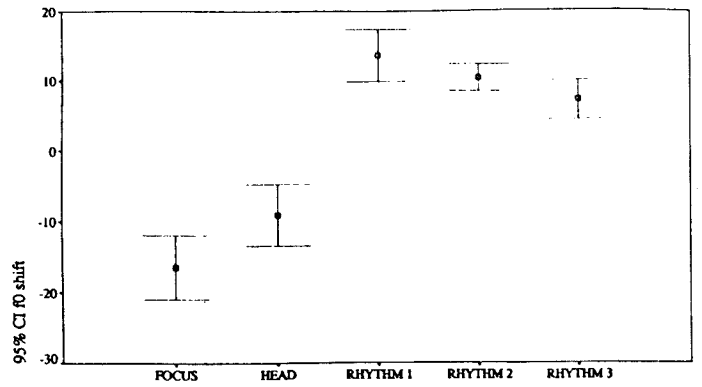

Figure 3. An error plot with 95% confidence interval of the mean relative shift in peak fundamental frequency in the shift words for each of the 5 contexts (in Hertz).