

# EFFECT OF DURATION, INTENSITY, AND F0 ALTERNATIONS ON RHYTHMIC GROUPING

Jessica F. Hay & Randy L. Diehl  
*University of Texas, Austin, Texas, USA*

## ABSTRACT

Stress patterns in language play a large role in how we perceive speech. Stress is assigned on the basis of duration, intensity, and F0. This study investigated the role that duration, intensity, and F0 play in rhythmic grouping. Subjects were presented with sequences of alternating squarewave segments that were varied on one of the aforementioned dimensions. They were asked to indicate whether they heard a trochaic or iambic rhythmic pattern. Results indicate that duration-varying sequences are grouped iambically, whereas intensity- and F0-varying sequences are grouped trochaically. Results replicate and extend previous research and may suggest a need for a reformulation of the Iambic/Trochaic Law.

## 1. INTRODUCTION

The perception of rhythm in English depends on the sequencing of stressed and unstressed syllables and how these are grouped by the listener into metrical units. An important question is how the perception of rhythm is effected by variation in three main correlates of stress: duration, intensity, and F0. Previous research [1, 4, 6] suggests that when subjects are presented with sound segments that alternate in intensity, a natural two-beat trochaic (strong-weak) rhythm is formed. In contrast, when alternating sound segments differ in duration, listeners tend to perceive the rhythm as iambic (weak-strong). Together, these two findings have formed the basis for the Iambic/Trochaic Law, which Hayes [2] states as follows:

- a. Elements contrasting in intensity naturally form groupings with initial prominence.
- b. Elements contrasting in duration naturally form groupings with final prominence. (p. 80)

Owing to certain methodological limitations, the research that has led to the formulation of the Iambic/Trochaic Law is in need of replication. The following vague generalizations made by Woodrow [6] illustrate this point:

[W]ith equal temporal spacing, and not too fast a rate, and every second sound louder than the others, the series of sounds tends to be heard in groupings of two, with the louder sound beginning the group. (p.1232)

The following study was conducted in order to replicate and extend previous findings on the role of duration, intensity, and F0 on the perception of rhythmic units. F0 was included as an additional variable because it is known to affect perception of stress. In each condition, sound segments were presented with equal inter-segment intervals and the segments themselves were varied in one of the three dimensions - duration, intensity, or F0. Subjects were asked to make judgments as to whether the sequences consisted of segment pairs having a strong-weak (trochaic) pattern or a weak-strong (iambic) pattern.

All stimuli in this study consisted of sequences of squarewave segments. The use of nonspeech stimuli is important for determining whether perceived rhythmic grouping in speech is a purely auditory phenomenon, or whether language itself plays a role. An additional study is currently underway in which the same stimulus variables are manipulated in synthetically produced CV sequences to determine whether the grouping principles carry over into the speech domain.

## 2. METHODS

### 2.1. Subjects

Twenty-four native English speakers participated in the study. All reported normal hearing. Subjects were recruited from the introductory psychology subject pool at the University of Texas at Austin, and were given credit toward partial fulfillment of course requirements. Fourteen additional subjects participated in a control condition.

### 2.2. Stimuli

The stimuli were created on a PC using an in-house waveform generating and editing program, Wavax, and consisted of sequences of square-wave segments, which alternated either in duration, intensity, or F0. The stimuli formed three sets. Within each set one of the three stimulus dimensions was manipulated while the other two were held constant. In the duration set, the alternating squarewave segments varied in duration while intensity and F0 were held constant. In the intensity set, the segments alternated in intensity while the while F0 and duration were held constant. In the F0 set, the segments alternated in F0 while duration and intensity were held constant. The silent interval between adjacent squarewave segments was fixed at 200 ms in all three stimulus sets. A control stimulus sequence consisted of squarewave segments that were constant in duration, intensity, and F0, and where the segments were again separated by 200 ms silent intervals. Table 1 shows the characteristics of all the stimulus sets plus the control sequence.

In pilot studies we noticed a strong tendency for the rhythmic judgments to be based on the pattern of the first two segments of the sequence. In order to reduce this effect we took steps to mask the onset on the sequence. Squarewave intensity was gradually ramped upward from zero over the first five seconds of the eleven second sequence. During this same five-second interval masking noise was presented with a gradually diminishing intensity.

Set, Sequence	Duration (ms)	Intensity (dB SPL)	F0 (Hz)
Duration, 1	100 & 200	56	100
Duration, 2	150 & 250	56	100
Control*	200	56	100

Intensity, 1	200	50 & 56	100
Intensity, 2	200	50 & 62	100

F0, 1	200	56	100 & 125
F0, 2	200	56	100 & 150
F0, 3	200	56	100 & 200

Table 1 Attributes for sequences in all stimulus sets.

\*When the control stimulus sequence is randomized within the set of duration-varying stimuli it is referred to as Control (D).

When it is presented in isolation in the control condition it is referred to as Control (I)

### 2.3. Procedure

Subjects were instructed to listen for the rhythm of each stimulus sequence and to indicate, by pressing a labeled button, whether the rhythm consisted of a strong sound followed by a weak sound, or alternatively, a weak sound followed by a strong sound. Segment sequences were presented five times each in both possible orders of onset and were randomized within each stimulus set. The order of presentation of stimulus sets was counterbalanced across subjects. At the beginning of each new stimulus set, subjects were given practice trials with each 11-sec sequence presented once in both orders. Randomized within the duration set subjects were presented with the control stimulus sequence. When judged in the context of the duration set, the control sequence is referred to as Control (D). Excluding practice trials, subjects hear a total of 25 duration-varying, 20 intensity-varying, and 30 F0-varying segment sequences.

Because it was unclear how subjects would assign strength weightings to F0 values, subjects were asked to perform the additional task of judging whether the high or low pitched segment of the F0-varying stimuli was the first sound of the rhythmic unit.

All subjects listened to the stimuli through headphones while seated at a response station in a double-walled sound-attenuated chamber.

Subjects in the control condition were presented with the control stimulus sequence ten times in isolation. In this condition the control stimulus is referred to as Control (I). They received two practice trials. All other procedures were identical to those described above.

## 3. RESULTS

One-way repeated measures ANOVAs were conducted for each stimulus set. There was no significant effect of order of presentation for the alternating squarewave segments (All  $F$ 's < .5,  $p > .5$ ). This suggests that the onset masking procedure was effective in eliminating the onset effect described above.

For the duration set (see Figure 1), when Control (D) was omitted from the analysis, there was no significant difference between Sequence 1 and 2 [ $F(3,69) = .53, p > .6$ ]. When Control (D) was included in the analysis, it yielded more trochaic

responses than either Sequence 1 or 2 [ $F(1,23) = 21.34, p < .000$ ]. This is in contrast to results found when the two duration-varying sequences were compared to Control (I) in the control condition. Given that the duration-varying sequences did not differ significantly, responses were averaged across sequences. An independent-samples t-test failed to yield a significant difference between the duration-varying sequences and Control (I) [ $t(36) = -.43, p > .6$ ].

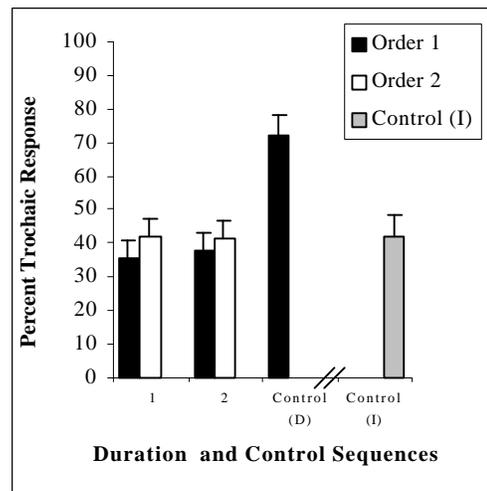


Figure 1. Percent trochaic response for duration-varying and control sequences. For each sequence, intensity and F0 are constant.

Sequence 1 has segments with durations of 100 ms and 200 ms.

Sequence 2 has segments with durations of 150 ms and 250 ms.

Control (D) and Control (I) have 200 ms segments.

A repeated measures ANOVA performed on the grouping of intensity-varying sequences showed that there was a significant difference between Sequence 1 (alternating segments differed by 6 dB) and 2 (alternating segments differed by 12 dB) [ $F(1,23) = 11.83, p < .002$ ] (see Figure 2). Sequence 2 was grouped trochaically more often (71.7%) than was Sequence 1 (55.9%). Whereas the grouping of Sequence 1 did not differ from Control (I) [ $t(36) = 1.80, p > .08$ ], Sequence 2 was grouped trochaically significantly more often than Control (I) (42.1%) [ $t(36) = 3.75, p < .001$ ].

The analysis of the effects of F0 on rhythmic grouping showed that there were no significant differences between Sequences 1 (100 Hz and 125 Hz), 2 (100 Hz and 150 Hz), and 3 (100 Hz and 200 Hz) [ $F(5,115) = .48, p > .7$ ] (see Figure 3). An independent sample t-test showed that the average response across the three sequences was significantly different from that of Control (I) [ $t(36) = 2.60, p < .013$ ].

When subjects were presented with F0-alternating sequences but instead were asked to judge whether the rhythmic group started with a high or a low pitch, the pattern of results changed considerably (see Figure 4). A two-way, repeated measures ANOVA showed that there was a significant effect of order of presentation for the alternating F0 segments [ $F(1,23) = 9.89, p < .005$ ]. There was also a significant difference between responses to Sequences 1, 2, and 3 [ $F(2,46) = 5.92, p < .005$ ]; as the difference in F0 increased, subjects were more likely to judge the rhythm as starting with the high F0.

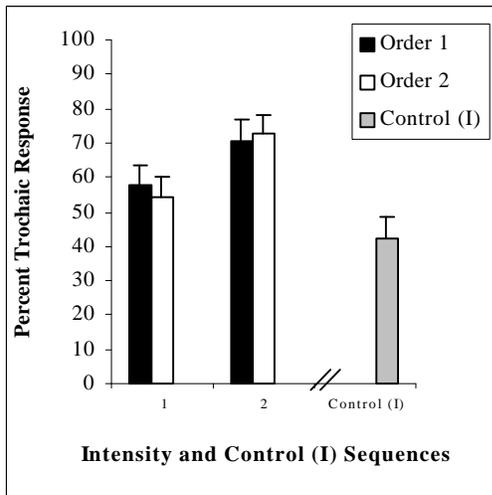


Figure 2. Percent trochaic response for intensity-varying and Control (I) sequences. For each sequence, duration and F0 are constant. Sequence 1 has segments with intensities of 50 dB and 56 dB. Sequence 2 has segments with intensities of 50 dB and 62 dB. The intensity of Control (I) is 56 dB.

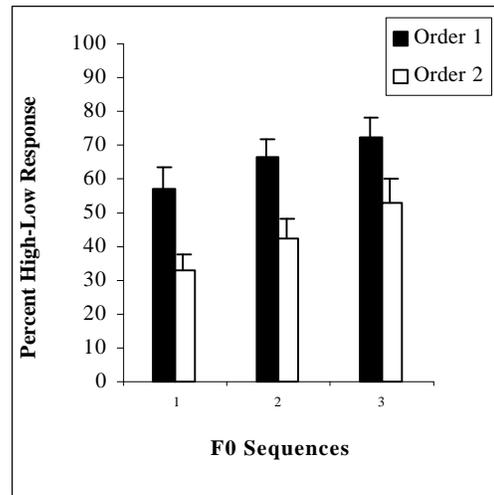


Figure 4. Percent high-low response for F0 sequences. For each sequence, duration and intensity are constant. Sequence 1 has segments with F0s of 100 Hz and 125 Hz. Sequence 2 has segments with F0s of 100 Hz and 150 Hz. Sequence 3 has segments of 100 Hz and 200 Hz

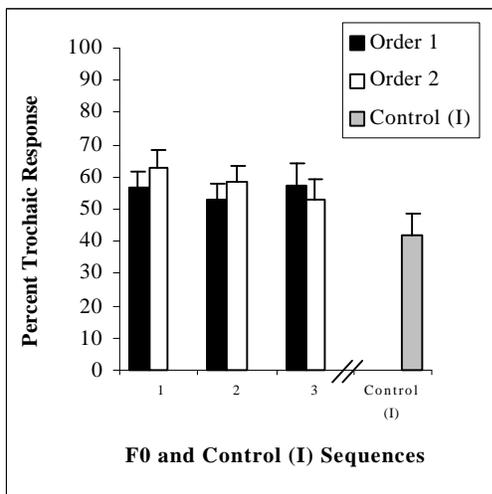


Figure 3. Percent trochaic response for F0-varying and Control (I) sequences. For each sequence, duration and intensity are constant. Sequence 1 has segments with F0s of 100 Hz and 125 Hz. Sequence 2 has segments with F0s of 100 Hz and 150 Hz. Sequence 3 has segments of 100 Hz and 200 Hz. The F0 of Control (I) is 100 Hz.

#### 4. DISCUSSION

The Iambic/Trochaic Law was partially supported by the results of this study. In particular, the sequences varying in intensity tended to be grouped trochaically and this tendency increased with the magnitude of the intensity difference. At first glance, the comparison between the duration varying sequences and Control (D) also appears to be consistent with the Iambic/Trochaic Law. This result, however, seems to be an artifact of presenting the control stimulus sequence within the context of sequences with unequal durations. When the control sequence was presented to subjects in isolation (Control (I)), it yielded a tendency toward iambic grouping that did not differ from the grouping of duration varying sequences. This result is surprising as it seems to indicate that varying the duration of alternating segments does not afford us anything over and above our natural tendency for grouping the control stimulus sequences.

Results on the effect of varying F0 showed some consistencies and some inconsistencies with earlier research [3, 4, 5, 6]. Woodrow [5] claimed that F0 itself has no predictable effect on rhythmic grouping. He also claimed, however, that sequences with segments of equal duration are grouped trochaically, and as such would predict a trochaic pattern for equal-duration, F0-varying sequences. Rice [3], on the other hand, found that there was a predictable effect of F0 on grouping pattern. His results indicated that F0 alternations result in an iambic rhythmic grouping. Like Rice, we found that F0 did have a predictable effect on rhythmic grouping. Unlike Rice, but like Woodrow, we also found that F0-varying sequences were grouped trochaically. Rice's F0 differences were quite small in comparison to ours - this might account for the discrepancy between our results.

Although the pattern of results observed for the F0-varying sequences indicates a trochaic grouping, how subjects assign strength to F0 is not known. If subjects were consistently assigning 'strong' to either the high F0 or to the low F0, the results for the high-low judgment should have been identical to those for the strong-weak judgment. We found that this was not the case. It appears that subjects process the same sequences differently when they are asked

to make different judgments. This raises the possibility that rhythmic grouping is not simply an auditory phenomenon, but instead involves cognitive factors.

Preliminary results from our second experiment using synthetically produced CV syllables seem to indicate that subjects process duration, intensity, and F0 in speech in much the same way that they process them in nonspeech.

## 5. CONCLUSIONS

Results from our experiment replicated and extended previous research. We found that alternating segments of unequal duration are grouped iambically. When duration is kept constant and intensity is varied, a trochaic rhythm emerges. The same is true when F0 is varied.

As far as extending previous research, this experiment is the first to have an adequate control condition. Like other researchers, we found that when sequences with segments of equal duration were presented amongst duration-varying sequences, they were perceived trochaically. Results from the control condition demonstrated that this is not an accurate description of how rhythmic grouping occurs when the control sequence is presented in isolation. This has some important implications for the Iambic/Trochaic Law and may ultimately necessitate its reformulation.

In future research we will investigate how the three correlates of stress interact when they are pitted against one another. We will also look at what effects the characteristic rhythmic patterns of specific languages have on perceived rhythmic grouping.

## ACKNOWLEDGMENTS

This work was supported by research grant number 5 R01 Dc00427-10, -11 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, to the second author.

## REFERENCES

- [1] Bolton, T. L. 1894. Rhythm. *American Journal of Psychology*, 6, 145-238.
- [2] Hayes, B. 1995. Metrical stress theory: Principles and case studies. Chicago: The University of Chicago Press.
- [3] Rice, C. 1992. Binariness and ternariness in Metrical Theory: Parametric extensions. Doctoral dissertation, University of Texas, Austin.
- [4] Woodrow, H. 1909. A quantitative study of rhythm. *Archives of Psychology*, 14, 1-66.
- [5] Woodrow, H. 1911. Role of pitch in rhythm. *Psychological Review*, 18, 54-77.
- [6] Woodrow, H. 1951. Time perception. In S. S. Stevens (ed.), *Handbook of Experimental Psychology* (pp. 1224-1236). New York: Wiley.