

Visual Rhythms: Dynamic Text Display for Learning to Read a Second Language

A method is described in which sentences are presented in a dynamic visual display. A television monitor is used to present simultaneously the visual and auditory versions of a sentence, with each of its successive visual and auditory syllables yoked in parallel; the onset of each visual syllable is synchronized with the onset of each syllable as it is heard through the auditory channel. The result is a sentence which "grows" left-to-right across the screen, one syllable at a time, in "visual rhythm." In an experiment, the subjects were three groups of secondary-school students learning Spanish as a second language. In training sessions, the rhythmic group saw the sentences in "visual rhythm," the unrhythmic group saw the same sentences but in static visual display, and the control group had no exposure to either visual-auditory display. Before and after training, all groups provided pre-test and post-test measures of oral reading fluency. The dependent measure was pre-test to post-test relative change in judged reading fluency. The results favored the rhythmic group.

Introduction

This paper proposes a novel method for visual presentation of sentences and other language materials as a way of facilitating the teaching of reading. The essential idea is that a sentence is printed in a dynamic display that changes through time. The method makes use of a TV monitor to present simultaneously the visual and auditory versions of a sentence, with each of its successive visual and auditory syllables yoked in parallel. This is accomplished by synchronizing the visual onset of each syllable with the onset of each syllable as it is heard through the auditory channel (the syllable rather than the word is the more appropriate temporal unit of speech). The result is a sentence which "grows" left-to-right across the screen, one syllable at a time, in "visual rhythm." To illustrate, say the sentence, "The boy is in the boat," and imagine that each syllable (the, boy, is, . . .) appears on the screen just as you speak it, so that the complete sentence is on display when the last syllable is spoken. The sentence may be presented visually either by itself or in combination with the auditory channel. The display functions as a link between a sentence as it is displayed visually in space and as it is displayed auditorily in time. It pre-

sents dynamic information about language in the visual mode, and it displays this dynamic information in what is probably the simplest and more relevant form for most purposes, readable text.

Our first experiment using the method (Martin and Meltzer, 1976) was an attempt to assess the reactions of young readers to the novel visual display, and to assess its possible effects on samples of their oral reading. The subjects were children in a summer remedial reading program who were to be second, third, or fourth graders in the fall. They were divided into two training groups. In several training sessions, the Rhythmic group was exposed to sentences appearing syllable-by-syllable in "visual rhythm" as described above. In training sessions for the Unrhythmic group, the same sentences were presented in a static display, that is, the complete sentence appeared all at once rather than syllable-by-syllable. In either condition the sentence was presented sometimes on both visual and auditory channels, sometimes on the visual channel alone, and on some trials the child attempted to read the sentence out loud. Each child had three training sessions of approximately 10 minutes spaced over a two-week period. At the beginning of the first training session, the child read aloud several sentences which were tape-recorded, and the same sentences were read aloud by each child and tape-recorded during the last session; together these sentences provided a pre-test and post-test measure of reading fluency for each child.

After training sessions were completed, each of the recorded pre-test and post-test readings of a given sentence by a given child were paired together, but in random order with respect to whether the pre-test or post-test reading was the first member of the pair. Finally, all utterance pairs were assembled on a new audiotape. Six judges listened to all utterance pairs and indicated which item of the pair sounded more fluent. The results of the experiment were in favor of the Rhythmic group. A higher percentage of post-test readings of a pair were judged to be more fluent in the case of the Rhythmic group than in the case of the Unrhythmic group.

The results suggested that the method might have some promise, but much more research, of course, was needed. Of the many alternatives which might be taken as the next step, we tried the method with another relevant and accessible population, second-language learners. In the experiment reported in this paper, the subjects were secondary-school girls who were taking a course in beginning Spanish, and as in the earlier experiment, the effects of training sessions were assessed by changes in pre-test and post-test oral reading fluency. There was a Rhythmic group and an Unrhythmic group differing in terms of intervening audiovisual training materials, and an additional Control group who provided pre-test and post-test readings without intervening experimental training with the visual-auditory display.

Method

Materials

Training sentences. Materials for the 39 sentences used in the experiment were based on the textbook used by the subjects for their course in beginning Spanish, and were chosen to supplement their classroom work. Sentences were selected or adapted from the text using the following criteria: (a) they contained less than 37 characters and spaces and (b) they contained only English punctuation and characters. These criteria were used because of limitations of the computer screen to be used in preparing the training materials. Two examples of sentences are: *Juan pone la ropa en la maleta. Alvedor de la casa hay una finca.*

The sentences, which ranged from 4 to 16 syllables in length, were read with normal intonation and rate by a native female Spanish (Cuban) speaker, who was the teacher of the Spanish class. Sentences were recorded in a sound-treated room on an Ampex AG-500 tape deck.

Preparation of training materials. The time-voltage waveform of each sentence was digitized and displayed on a PDP-12 computer screen in an oscilloscope-type display so that syllable onsets could be measured. Using this timing information, the rhythmic version of the sentences could be prepared. For this version a TV display was generated by the computer such that each syllable appeared visually

on the screen in synchrony with the auditory syllable. Each syllable appeared to the right of the preceding syllables until the entire sentence was displayed, after which the sentence remained on display for an additional 500 msec. A space appeared between each letter for greater clarity; triple spaces appeared between words.

Each rhythmic sentence had an unrhythmic counterpart. In this version the only difference was that all syllables in the entire sentence appeared on the screen simultaneously² and remained on the screen for the same length of time as the total time of the rhythmic sentence counterpart (visual first syllable onset to visual sentence offset). This total time ranged from about 1.2 to 3.0 sec.

Two parallel video tapes were constructed, one containing the rhythmic and the other containing the unrhythmic sentence versions. Each sentence was repeated five times sequentially on the tape with about 5 sec between repetitions and 6 sec between each set of 5 repetitions.

Pre-test and Post-test reading materials. These materials consisted of 16 sentences ranging from 7 to 24 syllables in length. Half of these were selected from the sentences displayed during training sessions (old), half were new. The new sentences contained familiar vocabulary but were otherwise unrelated to the training materials. The old and new sentences were mixed together and typed on a single sheet of paper.

Subjects and Procedure

The subjects were 45 girls enrolled in the second semester of a beginning Spanish course at their parochial secondary school, of which 14, 15, and 16 were in the Rhythmic, Unrhythmic, and Control groups, respectively. Seventeen additional girls participated in some portion of the experiment but were eliminated because they missed a training session or one of the reading fluency tests. Assignment of the subjects into the three groups was determined by scheduling convenience, but there were no significant differences in mean grades for the first semester of Spanish, and their teacher regarded the groups as alike in ability. All the girls were 9th or 10th graders with the exception of one girl who was in the 12th grade.

Prior to the training sessions, each subject was taken individually into a classroom and asked to read the 16 test sentences into a microphone for tape recording; these were the pre-test readings for the experiment. Each subject participated in five training sessions, each of which was 15 to 20 minutes long, in groups of from 5 to 16 subjects. The training sessions were held two times a week during regular Spanish class periods and spanned a 2-1/2 week interval.

The procedure during training session 1 for subjects in the Rhythmic and Unrhythmic groups was as follows. First the subjects arranged their chairs around the TV monitor (19" screen) so that everyone in the group could clearly see and hear the sentence. Then each

sentence was presented for five repetitions and responded to (according to their instructions) as follows. On the first two repetitions they watched and listened. On the third and fourth repetitions, they attempted to say the sentence aloud in time with the speaker. On the fifth repetition they again watched and listened. This routine was repeated for 21 different sentences.

The instructions were changed slightly during training sessions 2 and 3 since the sentences had been used during the first session and were now familiar. On the first repetition of each sentence the subjects watched and listened; on the next four repetitions they spoke the sentences aloud. During training sessions 4 and 5 a different set of 18 sentences was used. The instructions for training session 4 were like those of training session 1, since the 18 sentences were unfamiliar. The instructions for training session 5, after the sentences had become familiar, were like those of training session 2.

The Control group had no training sessions with the visual-auditory display. Instead, all of the materials from the old sentences in various contexts were made familiar to this group as part of their normal classwork routine. The results from this group provide a baseline for comparing changes over the time of the experiment.

After the completion of all training sessions, the 16 test sentences were read by each subject and tape-recorded for use as a post-test measure under identical conditions with the pre-test reading.

Listener Judgments of Reading Fluency

To reduce the size of the judgment task for the judges, only half of the 16 pre-test and post-test sentence pairs per subject were used. Four of these were old sentences, i.e., used during training; four were new sentences, ranging from 7 to 14 syllables. For each of the 45 subjects, each of the eight sentences recorded during the pre-test was paired with the identical sentence recorded during post-test, making a total of 358 pairs (360 minus 2 pairs from one subject which were missing because of inadvertent post-test recording mishap). Half of the 358 pairs were placed in the pre-test and post-test order in which they had been read, half were the reverse. All pairs were assembled on a new audiotape with the following constraints. Pairs were arranged in blocks of 45, of which approximately half (either 22 or 23) were reversed pairs. Each block contained one pair read by each of the 45 subjects. Each of the eight different test sentences read by at least four subjects was included in each block.

There were four judges, of which two were college-level Spanish instructors and two were college students whose native language was Spanish. The judges were informed of the randomized nature of the tape and were asked to make the fluency judgments on the basis of how smoothly sentences were spoken as a whole, and not on the basis of the pronunciation of individual words. They followed a typed manuscript of the sentences

while listening to the sentence pairs and judged which utterance of the pair was more fluent. They listened to each pair as often as necessary. After choosing one of the pair as more fluent than the other, each judge then gave a confidence rating for that choice on a five-point scale.

Results

A judgment that the post-test reading of a given sentence pair by a given subject is more fluent indicates improvement over the time of the experiment, and percentages and mean confidence ratings across subjects permit a comparison of relative improvement among the three training groups in the reading of both old and new sentences. Table I gives the summary data. The results are quite similar for direction of judgment (percentages) and mean confidence ratings. The result of interest is the relatively greater improvement by the Rhythmic group in reading aloud the new sentences during the post-test.

The percentages were tested by summing the number of pair judgments in favor of the post-test reading for each subject and each judge. With four sentence pairs this value could range from zero to four. These sums were submitted to analysis of variance with sentence type (old/new) as a within-subjects factor and training condition (Rhythmic/Unrhythmic/Control) and judge as between-subjects factors. The following results were significant: sentence type,

Table I
 Percent Judgments and Mean Confidence Ratings
 (Ratings in Parentheses)

Sentences	Group		Control	All
	Rhythmic	Unrhythmic		
Old	79.5% (2.58)	80.0% (2.54)	64.5% (1.36)	74.7% (2.16)
New	73.1% (1.71)	59.2% (0.80)	63.7% (0.97)	65.3% (1.16)
All	76.3% (2.15)	69.6% (1.67)	64.1% (1.17)	70.0% (1.66)

$F(1,168)=16.80, p<.01$; training condition, $F(2,168)=5.85, p<.01$; and the interaction of sentence type and training condition, $F(2,168)=6.16, p<.01$.

Mean confidence ratings were tested by assigning negative sign to judgments in favor of pre-test rather than post-test readings and averaging across the four sentences; these values could range from -5 to $+5$ with positive values in favor of the post-test readings. Analysis of variance yielded the following significant factors: sentence type, $F(1,168)=27.30, p<.01$; training condition, $F(2,168)=7.93, p<.01$; judge, $F(3,168)=4.34, p<.01$; and interaction between sentence type and training condition, $F(2,168)=4.14, p<.025$.

Separate analyses of the Rhythmic vs Unrhythmic conditions alone yielded significant differences also, as did an analysis using sentences rather than subjects as the sampling unit.

Discussion

The relatively greater improvement by Rhythmic and Unrhythmic groups over the Control group can be attributed to practice in pronouncing sentences aloud with the help of the visual-auditory display, and relatively greater improvement (Rhythmic and Unrhythmic groups) in old sentences over new sentences can be attributed to practice with these particular sentences. The difference favoring the Rhythmic group over the Unrhythmic group in the case of new sentences suggests greater transfer to fresh sentences after practice with the rhythmic display. This result generalizes the earlier result (Martin and Meltzer, 1976) in favor of the visual rhythm over the static display.

Speculating on the reasons for differential transfer effects from the rhythmic and static display, little can be said on the basis of observations of individual behavior during the training sessions since the subjects were run in groups. But it is plausible to suggest that since the rhythmic visual display duplicates timing information in the auditory channel, it helped the subject to know when she was successfully shadowing the auditory channel and when she was lagging behind. The dynamic display also tends to convey a sense of rapid speaking rate, and this is another reason why it may be an effective pacer for oral reading practice, as it appears to be. Of course these considerations predict greater improvement for not only new but old

Sounds and symbols are paired dynamically so that many-to-one and one-to-many correspondences between visual and auditory modes may be noted. In pluralization, for instance, the "s" in "caps" and in "cabs" is seen as the same in both cases in visual display, but is heard on the auditory channel as /s/ in the former but /z/ in the latter. When "cake" is pluralized we see and hear the syllable ending in /s/. But when "horse" is pluralized, we hear it as two syllables, the latter of which is /lz/, and we also simultaneously see it as two syllables since they appear on the screen temporally separated. Examples could be multiplied. While some aspects of orthography are handled more naturally than others, the display might facilitate the concurrent part and whole learning in a rather efficient way. The display is quite flexible in its technical aspects; its usefulness can be evaluated in other ways than by oral fluency, and it can be used with various populations, including the deaf who are known to have particularly difficult problems in learning to read (e.g., Odom, 1974). It is possible that any population with reading difficulties might profit from experience with the dynamic display.

In conclusion, our work with the visual rhythms display is an attempt to bring some of our ideas about speech perception (Martin, 1972; Meltzer, Martin, Mills, Imhoff and Zohar, 1976) to bear on the reading process by focussing on dynamic aspects of the stimulus. It is not clear that this work can be used to choose among

sentences also, in a comparison between Rhythmic and Unrhythmic groups. But old sentences were seen and spoken repeatedly (10-15 times each) during training sessions by both groups. This much familiarity with the sentences could easily provide a ceiling on the advantages of a rhythmic over a static display. Additionally, one might even hypothesize a negative transfer component to old sentences in the Rhythmic group which could have partially cancelled positive transfer effects, since for this group these sentences were in dynamic display during training but static display during post-test.

The studies we have done using oral fluency tend to emphasize holistic aspects of learning to read, but it is possible that the visual rhythms display will be useful for the learning of part aspects also, for instance, the learning of sound-symbol correspondences. Learning to process written language is often assumed to be more difficult than learning to process spoken language (Kavanagh and Mattingly, 1972). One likely bottleneck is that the beginning reader must first learn to decode the parts (letters, words) of a sentence rapidly, or else the earlier items in a sentence are lost from memory before they can be integrated (Gough, 1972). In other words, parts must be mastered before wholes can be grasped, in learning to read a sentence. A potential advantage of the visual rhythms display is that part and whole aspects of learning to read may proceed concurrently. Consider the learning of sound-symbol correspondences.

competing models of the skilled reader (Massaro, 1975), or that it can illuminate much of what takes place during any of the various kinds of skilled reading (Gibson and Levin, 1975). But the outlook for the study of learning to read seems more promising. As Massaro characterized Gough's (1972) view of the situation: "When a child begins

to learn to read, he already knows how to listen. If somehow the printed symbols on the page could be made to speak, all that the young reader would have to do is listen" (Massaro, 1975, p. 260). We can almost agree with this, without having to believe that it comes anywhere near to the whole story.

1 This research was supported in part by a grant from the Army Research Institute for the Behavioral and Social Sciences and by a grant from the Bio-Medical Sciences Support Committee, University of Maryland, to the Center for Language and Cognition. Computing services were provided by the Computer Science Center of the University of Maryland. Thanks are expressed to Maria Torres, Clara Porcel, and the rest of the faculty of Regina High School, Hyattsville, Maryland, for their advice and their help in providing subjects for the experiment. Robin Jeffries was the research assistant for this project. Requests for reprints should be sent to James G. Martin, Psychology Department, University of Maryland, College Park, MD 20742.

2 This display was not quite simultaneous; computer hardware limitations required that each letter in the display appear sequentially, in this case at 8 msec intervals. Hence while letters within, say, a three-letter syllable appeared to be simultaneous in the Rhythmic version, the sequencing was slightly noticeable in the Unrhythmic version, since a sentence containing; e.g., 18 letters and spaces took 144 msec before the display was complete. In this case, the effects were as though the letters appeared across the screen in a "burst."

3 The criteria for judgments although vague are apparently not biased in favor of any particular group. However, future research should include separate judgments of timing, inflection, word pronunciation, etc.

References

- Gibson, E. J., & Levin, H.
The Psychology of Reading.
Cambridge, Mass.: MIT Press, 1975.
- Gough, P. B.
One second of reading. In J. F. Kavanagh and I. G. Mattingly (Eds.), *Language by ear and eye, the relationships between speech and reading*.
Cambridge: MIT Press, 1972.
- Kavanagh, J. F., & Mattingly, I. G. (Eds.),
Language by Ear and by Eye.
Cambridge: MIT Press, 1972.
- Martin, J. G.
Rhythmic (hierarchical) versus serial structure in speech and other behavior.
Psychological Review,
79, 487-509, 1972.
- Martin, J. G., & Meltzer, R. H.
Visual rhythms: Report on a method for facilitating the teaching of reading.
Journal of Reading Behavior,
1976, 8, 153-160.
- Massaro, D. W.
Primary and secondary recognition in reading. In D. Massaro (Ed.), *Understanding Language: An information-processing analysis of speech perception, reading and psycholinguistics*.
New York: Academic Press, 1975.
- Meltzer, R. H.; Martin, J. G.; Mills, C. B.; Imhoff, D., & Zohar, D.
Reaction time to temporally-displaced phoneme targets in continuous speech.
Journal of Experimental Psychology,
1976, 2, 277-290.
- Odom, P. B.
Some speculations concerning why the deaf cannot read well. Paper presented to SRCD Conference, New Orleans, August, 1974.