enhance timing accuracy, whereas concentration on anything else would probably interfere with it. By contrast, the behavioral correlates of drowsiness are much simpler.

The chief significance of the present study lies in the demonstration of a positive relationship between two continuous variables, one of which is physiological (alpha voltage) and the other psychological (accuracy of time perception). Presumably, this represents a way in which the quantitative physiological psychology of consciousness can be advanced beyond the limitations of arbitrarily defined "stages." While these data might tempt some to speculate about the "slowing of a neural clock," it would seem that the deterioration of perception and memory associated with a decrease in quality and quantity of alpha activity reported by Simon and Emmons (2) for conditions of drowsiness is sufficient to account for the errors in time perception observed in the present study (12).

JAMES ANLIKER
Anesthesia Laboratory, Harvard Medical School, Massachusetts General Hospital, Boston 14

References and Notes

1. H. Blake, R. W. Gerard, N. Kleitman, *J. Neurophysiol.* **2**, 48 (1939); A. L. Loomis, E. N. Harvey, G. Hobart, *J. Exptl. Psychol.* **21**, 127 (1937).
2. C. W. Simon and W. H. Emmons, *Science* **124**, 1066 (1956).
3. Since localization was not the primary goal of this experimental study, the transoccipital linkage (between right and left occipital electrodes) was selected.
4. Electroencephalograph, model III-D, Grass Instrument Co., Quincy, Mass.
5. H. C. Becker, W. A. Mickle, R. G. Heath, *Electroencephalog. Clin. Neurophysiol.* **10**, 731 (1958).
6. Operational amplifier, model K2-W, George A. Philbrick Researches, Boston, Mass.
7. Analogue-to-discrete data conversion: a circuit provided for the momentary operation of a relay whenever an arbitrarily selected level of voltage integration was reached. The operation of the relay provided an output pulse and re-set the integrator to zero voltage.
8. Harvard cumulative recorder, Ralph Gerbrands Co., Arlington, Mass.
9. Only the initial key depression produced a pulse in the control circuit, so that holding the key down had no further effect. In order to make another response the subject had to release the key and press it again.
10. Noise generator, model 901; electronic switch, model 829B; recording attenuator, model E3262A; all by the Grason-Stadler Co., W. Concord, Mass.
11. W. Dement and N. Kleitman, *J. Exptl. Psychol.* **53**, 339 (1957).
12. Supported in part by research grant A.F. 49-(638)-98 from the Air Research and Development Command, Department of the Air Force, and in part by PHS research grants MY-2971 and MY-00987 from the National Institute of Mental Health. The constant support and helpful criticism of Dr. Henry K. Beecher, in whose laboratory this research was conducted, are gratefully acknowledged. Much guidance on neurophysiological matters was graciously provided by Dr. Mary A. B. Brazier. I am also indebted to Melvin Meister for technical assistance.

29 April 1963

Redundancy in Children's Texts

Abstract. *The sequential constraints of pairs of letters were computed for a series of graded readers. Information in single letters is about the same for all texts. Redundancy decreases with increasing grade in a regular way, while mean word length increases. A third reader has about the same redundancy as simple adult text. The constraints in a first reader are considerable, whereas those in a fifth reader approach those in average adult text.*

There is good evidence that letter redundancy is a powerful determinant of the learning of verbal materials (*1*) and of the reconstruction of mutilated texts (*2*). That such contextual constraints affect the learning of similar material by children has been shown also (*3*). The application of information theory to the sequential constraints among the letters has permitted quantification of the letter redundancy in texts of varying difficulty (*4*). Children's graded readers are written on the assumption that a small initial vocabulary, increasing within a reader and within a series, yields texts that become progressively more difficult. If this is true, then sequential constraints should decrease as the difficulty of the reader increases. As a first step toward quantification of language development in children, reading texts at levels 1, 2, 3, and 5 were compared with each other and with texts intended for adults.

We have measured letter redundancy by using a slight modification of Newman and Waugh's (*4*) method. Our modification gives more precise estimates because of more extensive tables of $\log_2$ and addition of the Miller-Madow (*5*) bias correction for sample size computed at each stage in a sequence. Contingency tables were constructed, each of which showed the number of times that a given symbol was followed at step n to the right by every possible symbol, where n takes on the values 1, 2 . . . , n. The number of symbols, a, was 28—the 26 letters of the English alphabet plus the space (for word length) and period (for sentence length). From these contingency tables, calculations were made of

$$H = -\sum_{i=1}^{a} p(i)\ \log_2 p(i) \qquad (1)$$

in which H is the amount of information in bits and $p(i)$ is the probability associated with each of the alternative symbols. This is just the information contained in letters according to their relative frequency of occurrence, which we write hereafter as $H(1)$. The amount of information found in pairs of adjacent items, i and j, in a sequence is $H(i, j)$, where

$$H(i, j) = -\sum_i\sum_j p(i, j)\log_2 p(i, j). \qquad (2)$$

If row and column selections were made independently, then a given table would contain twice the amount of information computed for single letters. The information added by the second letter of the pair when the first is known, $H_1(2)$, is:

$$H_1(2) = H(1, 2) - H(1) \qquad (3a)$$

or, more generally, the amount of information which is added by the nth letter of a sequence of n letters when the first is known, $H_1(n)$, is:

$$H_1(n) = H(1, n) - H(1). \qquad (3b)$$

McGill and Garner (*6*) define a quantity, T, the conditional information in a second letter that is dependent on the choice of a first letter

$$T(1, n) = 2H(1) - H(1, n). \qquad (4)$$

The information in the last letter of a sequence $F(n)$, may be approximated (neglecting interaction terms) as

$$F(n) = H(1) - \sum_{k=1}^{n} T(1,k). \qquad (5)$$

Finally, the relative sequential constraint in a sequence of n letters, C_n is given by

$$C_n = \frac{\sum_{k=2}^{n} T(1, k)}{H(1)}. \qquad (6)$$

Appropriate information measures were computed to determine the constraints between all pairs of letters at distances of the second letter from the first of 2, 3, . . . , 12, 30, 60, and 120.

Table 1 shows that constraint increases regularly from the *Atlantic Monthly* to the First Reader, with the Bible being about as constrained as the Third Reader, and the Fifth Reader approaching average adult text (*7*). This is true whether only the constraint of adjacent pairs is considered or whether the constraint between the first and last letters in sequences of 11 letters is considered. The relative rank ordering of C_1 and C_{11} contains but one reversal, the Bible, which is a peculiar text in terms of repetitiveness of words of low frequency in Modern English and also is represented by a small sample (*8*). $H(1)$ is about the same for all texts,

Table 1. Relative sequential constraint, C_2, is shown for a variety of texts. In some cases C_{11} is shown also. The last three columns give mean word length, letter sample size, and $H(1)$. The alphabet size used in computations on English letters was 28—26 letters plus space and period. The alphabet size for the Russian Bible was 35 letters. The alphabet size for the phonemic system of the urban Russian speaker was 42 phonemes.

Source	C_2	C_{11}*	Mean word length	Number of letters in sample	$H(1)$
Atlantic Monthly†	.184	.319	4.653	10,000	4.152
William James†	.192	.361	4.556	10,000	4.121
5th Reader	.213	.378	4.106	64,454	4.113
3rd Reader	.221	.414	3.997	55,027	4.131
Bible	.216	.429	4.014	10,601	4.049
2nd Reader	.233	.486	3.748	61,759	4.125
1st Reader	.241	.518	3.665	49,556	4.105
Russian Bible	.245	.472	5.296	10,000	4.549
Russian phonemes‡	.232			9,000	4.780

* The usual measure of percentage redundancy is given by multiplying C_{11} by 100. † Computed from data of E. B. Newman and N. C. Waugh, *Information & Control* **3**, 101 (1960). ‡ Computed from data of R. Jakobson, E. C. Cherry, M. Halle, *Language* **29**, 34 (1952).

approximately 4.1, the value which would be obtained for an optimally encoded alphabet of 18 letters. It should be noted that C_{11} and mean word length are perfectly and inversely correlated. This suggests that constraints are probably heavily determined by the size of the lexicon. The essence of the graded reader is to introduce new and larger words within a reader as well as between readers. This fact, taken together with the finding that $H(1)$'s are about equal for all texts, supports the notion that constraint lessens as lexicon size grows. However, recent analysis of a child's story book shows that even with relatively long mean word length (4.197 letters), repetition of words still yields high redundancy of (C_{11} = .446).

The extremely regular nature of the function C_n for successive values of n from 2 to 11 may be seen in Fig. 1. Note that different texts reach asymptote at different values and that no text shows any appreciable increase in constraint for n greater than 9. In fact, values of C_n for n = 12, 30, 60, and 120 are not sensibly different from C_9 for a given text.

Equal $H(1)$ implies identical distributions of frequency of use of letters. It is not implied, however, that the same letter be used with the same frequency in two different texts. The rank correlations of relative frequency of use of letters were computed for children's readers and general adult English texts. All the rank order intercorrelations were 0.95 or higher. Thus a stronger conclusion is in order. Not only are the distributions of frequency of use of letters essentially the same for children's texts as for adult texts, but it is also true that the rank order of frequency of use of a given letter is about the same for a child's text as for an adult text.

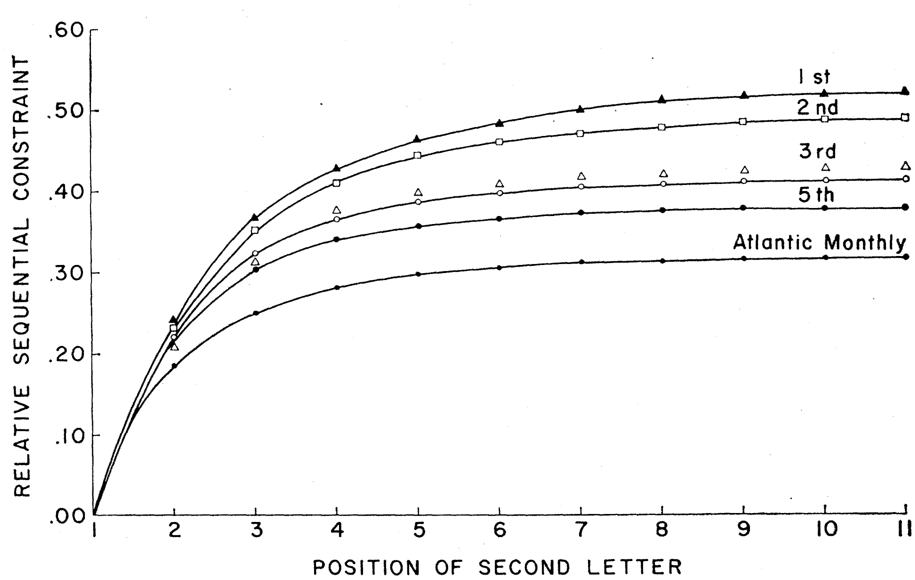

Fig. 1. Relative sequential constraint, C_k, of the letter in the k^{th} position. Four graded readers are shown. The *Atlantic Monthly* gives a lower bound for relatively complex texts. The English Bible is shown by open triangles.

This method of computing relative sequential redundancies may be used to quantify sequences of symbols of any kind. For example, a sample of text or of speech could be recoded as a string of grammatical units, or as a sequence of words, or as a string of phonemes. From the first, one could compute a measure of constraint which would be characteristic of syntactic style, from the second a measure of richness of vocabulary, and from the third a measure of redundancy of spoken and informal language. The last of these is important in making statistical comparisons of spoken and written language and in studying the early development of language in the child. There is some suggestive evidence to support the conjecture (*2*) that the relationship is quite close. For example, C_2 computed from a 9000 phoneme sequence of an urban Russian speaker (*10*) is 0.232. C_2 is 0.245 for the Russian version of the King James Bible (*4*). These two values are remarkably close. Their similarity is of the same order as that for the English Bible and *William James*, two samples of written language. Naturally, these comparisons are to be taken only as tentative, for the first comparison is between letters and phonemes while the second is between letters.

The redundancy of children's spontaneous language should be compared with that of their readers in order to ascertain whether the difficulty of the readers is indeed suited to the language of the child. After the values of C_k for the spontaneous language are known, it will be possible to construct stimulus materials with known constraints for use in language-learning experiments. The aim of such experiments is to understand the growth of patterning in children's language and the relationship of the optimal sequential redundancy, as measured in the learning experiments, to the sequential redundancy in their own speech and to that in language materials intended for them. It is important to use measures based on phonemic strings so that one can study the language of very young children several years before they are able to read or write. Thus language behavior may be investigated at a stage before and during the growth of constraints imposed by the natural language (*11*).

EDWARD C. CARTERETTE
MARGARET HUBBARD JONES
Department of Psychology, University of California, Los Angeles

References

1. A. D. Baddeley, paper presented before the British Exptl. Psychol. Society, Jan. 1962; L. M. Horowitz, *J. Exptl. Psychol.* **62**, 51 (1962); G. A. Miller, *ibid.* **56**, 484 (1958).
→ N. G. Burton and J. C. R. Licklider, *Am. J. Psychol.* **68**, 650 (1955); W. R. Garner and D. H. Carson, *Psychol. Repts.* **6**, 123 (1960); G. A. Miller and E. A. Friedman, *Information & Control* **1**, 38 (1957); C. E. Shannon, *Bell System Tech. J.* **30**, 50 (1951). See also the review by W. R. Garner, *Uncertainty and Structure as Psychological Concepts* (Wiley, New York, 1962).
3. R. S. Bogartz and E. C. Carterette, *J. Exptl. Psychol.*, in press; D. D. Kluppel, dissertation, U.C.L.A., 1962.
4. E. B. Newman and L. J. Gerstman, *J. Exptl. Psychol.* **44**, 114 (1952); E. B. Newman and N. C. Waugh, *Information & Control* **3**, 101 (1960); C. E. Shannon, *Bell System Tech. J.* **27**, 379 (1948).
5. G. A. Miller and W. C. Madow, Air Force Cambridge Res. Center, *AFCRC-TR-54-75* (1954).
6. W. R. Garner and W. J. McGill, *Psychometrika* **21**, 219 (1956).
7. Texts were the Ginn and Company basic readers: *On Cherry Street* (1st reader, 1957 revision); *We Are Neighbors* (2nd reader, level I, 1957 revision); *Finding New Neighbors* (3rd reader, level I, 1957 revision); *Trails to Treasure* (5th reader, 1957 revision). The number of different words in the vocabularies of the readers were 326 (1st), 552 (2nd), about 1088 (3rd), and 2792 (5th).
8. See R. Seibel, "*N*-Gram (sequences of letters *N* long) analysis of English for *N* from 1 through 12," Eastern Psychol. Assoc. Mtg., April 1962, for some effects of sample size.
9. G. Dewey, *Relativ* (sic) *Frequency of English Speech Sounds* (Harvard Univ. Press, Cambridge, Mass., 1923), p. 185.
10. R. Jakobson, E. C. Cherry, M. Halle, *Language* **29**, 34 (1952).
11. We thank Dr. Peter M. Neely, U.S. Public Health Service postdoctoral fellow in the School of Public Health, University of California, Los Angeles, for writing a general and flexible program for the I.B.M. 7090 computer. We are also indebted to the Western Data Processing Center, University of California, Los Angeles, for the use of computing facilities. Supported by U.S. Office of Education cooperative research project No. 1877.

15 April 1963

Ecdysone: Five Biologically Active Fractions from Bombyx

Abstract. *Five fractions of the growth and differentiation hormone, ecdysone, separated from extracts of* Bombyx *have been detected by bioassay. Three of these have not been described heretofore.*

Growth and metamorphosis of insects is controlled by the interaction of several hormones. Cholesterol has been identified in one of the biologically active extracts of brain from the silkworm that activate Dauer pupae (*1*). Farnesol and its oxidation product, farnesal, isolated from *Tenebrio* (*2*) have activity similar to that of the secretion of the corpora allata, neotenin or juvenile hormone (*2*). The growth and metamorphosis hormone, ecdysone, was originally isolated as α- and β-fractions, and crystallized (*3*); its chemical structure and biological activity have been the subject of further study (*4*). This hormone increases the rate of protein synthesis in mammalian cells (*5*). Puffs in the salivary chromosomes (*6*) were originally described in *Drosophila* (*7*) and the sequence of appearance delineated later (*8*). The relationship between the appearance of the puffs, hormonal secretion, and stage of metamorphosis has been studied by a number of investigators (*9*). We have confirmed (*10*) in *Drosophila* results obtained in *Chironomus* (*11*) in which the pattern of puffing in the salivary chromosomes is altered by ecdysone.

Ecdysone was extracted (*12*) from chrysalides of *Bombyx* obtained from Japan in ton lots. The method (*3*) was changed by utilizing *Calliphora* for bioassay to determine the distribution of the crude hormone between water and several immiscible organic solvents so that higher yields were obtained. Each bioassay was performed by injecting 10 μl of hormone into the posterior segment of each of 20 ligated *Calliphora* pupae with the anterior end pupated during the antecedent 24 hours. The degree of pupation of each individual was scored as 1.0, 0.75, 0.50, or 0.0 during the ensuing 24 hours and expressed as a percentage of the number surviving. After active crude extract was obtained, it was partitioned at first with a cyclohexane-butanol-water system. An ethyl acetate-water system later proved to be more satisfactory. By using a Craig counter-current machine with 200 transfers, active material was found, not only in the tubes expected for isolation of α- and β-ecdysone, but also farther along in the series of fractions. Careful bioassay revealed five peaks of activity separable not only by a threshold of 50 percent activity in the system of weighted bioassay but as low as 30 percent as well (Fig. 1). The same was true with the cyclohexane-butanol-water system.

Apparently there are five separate fractions that can be isolated from *Bombyx* at the stage of metamorphosis when the titer of crude hormone is highest (*13*). The γ-, δ-, and ε-fractions have not been noted previously. Theoretical *K* values (partition coefficients) for the compounds yielding this peak biological activity in the ethyl acetate-water system are 0.41, 0.96, 2.2, 6.2, and 50, respectively.

Separation of the hormone or hormones that bring about metamorphosis into individual components opens additional avenues for exploring the mechanics of growth and differentiation. Initial experiments indicate that crude material alters the pattern of growth of mammalian cells as well as accelerating the rate of protein synthesis in the cytoplasm (*14*); therefore study of the relationship between the action of invertebrate hormones and differentiation may be profitably extended to vertebrates. The results of studies in progress on possible differences between the effects of the fractions isolated should be of particular interest (*15*).

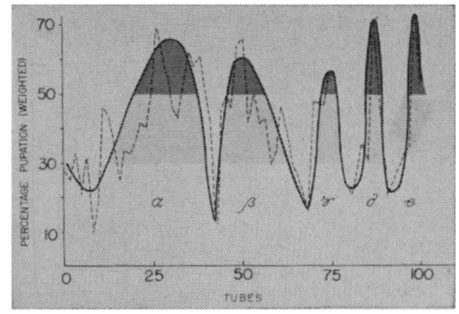

Fig. 1. Distribution of five fractions of the growth and metamorphosis hormone, ecdysone, detected by bioassay of fractions separated by counter-current distribution. Solid lines represent theoretical extrapolation of actual values (dotted lines). The figure depicts the 100-tube equivalent of a 200-tube run.

WALTER J. BURDETTE

Laboratory of Clinical Biology and *Department of Surgery, University of Utah College of Medicine, Salt Lake City*

MILON W. BULLOCK

Chemical Research and Development Laboratories, Agricultural Division, American Cyanamid Company, Princeton, New Jersey

References and Notes

1. M. Kobayashi, J. Kirimura, M. Saito, *Mushi* **36**, 85 (1962); *Nature* **195**, 515 (1962).
2. V. P. Schmialek, *Z. Naturforsch.* **16b**, 461 (1961); V. B. Wigglesworth, *J. Insect Physiol.* **7**, 73 (1961).
3. A. Butenandt and P. Karlson, *Z. Naturforsch.* **9b**, 389 (1954).
4. P. Karlson, *Angew. Chemie*, in press (1963).
5. W. J. Burdette and R. L. Coda, *Proc. Soc. Exptl. Biol. Med.* **112**, 216 (1963).
6. G. T. Rudkin, *Genetics* **40**, 593 (1955).
7. T. S. Painter, *ibid.* **20**, 301 (1935).
8. H. J. Becker, *Chromosoma* **10**, 654 (1959).
9. M. E. Breuer and C. Pavan, *ibid.* **7**, 371 (1955). W. Beerman, *Cold Spring Harbor Symp. Quant. Biol.* **21**, 217 (1956).
10. W. J. Burdette and R. S. Anderson, unpublished.
11. U. Clever and P. Karlson, *Exptl. Cell Res.* **20**, 623 (1960); U. Clever, *Chromosoma* **12**, 607 (1961).
12. The technical assistance of L. Paul, K. Nomura, R. Berndt, and G. Henkel is gratefully acknowledged.
13. W. J. Burdette, *Science* **135**, 432 (1962).
14. ———, *ibid.* **138**, 987 (1962); ——— and R. C. Richards, *Nature* **189**, 666 (1961).
15. Aided by grants from the National Institutes of Health, U.S. Department of Health, Education, and Welfare, and the American Cancer Society.

29 March 1963