

The Algorithmic Complexity of Neural Spike Trains Increases During Focal Seizures

P. E. Rapp,¹ I. D. Zimmerman,¹ E. P. Vining,¹ N. Cohen,¹ A. M. Albano,² and M. A. Jiménez-Montaño³

¹Department of Physiology, The Medical College of Pennsylvania, Philadelphia, Pennsylvania 19129, ²Department of Physics, Bryn Mawr College, Bryn Mawr, Pennsylvania 19010, and ³Department of Physics and Mathematics, Universidad de las Americas/Puebla, Cholula 72820, Puebla, Mexico

The interspike interval spike trains of spontaneously active cortical neurons can display nonrandom internal structure. The degree of nonrandom structure can be quantified and was found to decrease during focal epileptic seizures. Greater statistical discrimination between the two physiological conditions (normal vs seizure) was obtained with measurements of context-free grammar complexity than by measures of the distribution of the interspike intervals such as the mean interval, its standard deviation, skewness, or kurtosis. An examination of fixed epoch data sets showed that two factors contribute to the complexity: the firing rate and the internal structure of the spike train. However, calculations with randomly shuffled surrogates of the original data sets showed that the complexity is not completely determined by the firing rate. The sequence-sensitive structure of the spike train is a significant contributor.

By combining complexity measurements with statistically related surrogate data sets, it is possible to classify neurons according to the dynamical structure of their spike trains. This classification could not have been made on the basis of conventional distribution-determined measures. Computations with more sophisticated kinds of surrogate data show that the structure observed using complexity measures cannot be attributed to linearly correlated noise or to linearly correlated noise transformed by a static monotonic nonlinearity. The patterns in spike trains appear to reflect genuine nonlinear structure.

The limitations of these results are also discussed. The results presented in this article do not, of themselves, establish the presence of a fine-structure encoding of neural information.

[Key words: epilepsy, focal seizure, algorithmic complexity, surrogate data]

This investigation begins with the observation that neural messages are, in part, encoded in the interspike intervals, I_1 , I_2 , I_3 , In its most elementary form, the analysis of these data

begins with an examination of the statistical properties of the distribution formed by $\{I_i\}$ such as the mean, median, and SD. In more sophisticated studies additional properties of the distribution such as skewness and kurtosis may also be determined. The value of this analysis is well established. For example, Selz and Mandell (1992) report that differences in the higher moments of the interspike interval distribution can be used to discriminate between three kinds of biogenic amine neurons. These measures are, however, insensitive to the sequence of the spike trains. The spike train obtained by randomly shuffling the original data has the same distribution. Average values of the interspike interval and the SD, for example, are unchanged even though the internal structure of the original message has been lost.

Efforts to find sequence-sensitive patterns in neural spike trains have a long history (Perkel and Bullock, 1968). In a pioneering series of reports, Sherry and Klemm (1982a,b, 1984; Klemm and Sherry, 1982) compared three procedures for characterizing interspike interval data: measures of the probability density function, Markov order, and entropy. They demonstrated that calculations of Markov order and of entropy found evidence for structure in spike trains that was not found by conventional statistical measures of the interspike interval distribution. A well-known example of a sequence-sensitive measure is the method developed by Dayhoff (Dayhoff and Gerstein, 1983a,b; Dayhoff, 1984). This procedure searches for preferred words in spike trains. Using a modified variant of this method, Chen and Ku (1992) found differences in the structure and repetition frequency of favored patterns in spontaneous and evoked activity in the substantia nigra and in hypothalamic nuclei. Other procedures for analyzing the fine structure of neural spike trains have been published by Lestienne and Strehler (1987; Strehler and Lestienne, 1986), Legéndy and Salzman (1985), Abeles and Gerstein (1988), Vaadia et al. (1989), Villa and Abeles (1990), and Mandell and Selz (1993a,b). In this article we will apply a sequence-sensitive measure of complexity to records obtained from rat cortical neurons before and after the topical application of penicillin. This procedure and others related to it, notably the alumina focus, also have a long history. Single-unit recordings from epileptogenic foci in animals (e.g., Wyler et al., 1973; reviewed in Wyler and Ward, 1980) and in man (Calvin et al., 1973) have been analyzed by a variety of procedures including spectral analysis (Tepper and Mandell, 1987) and the burst index (Wyler et al., 1975; Schmidt et al., 1976).

In the course of this investigation we will address what we believe to be an essential question concerning the value of this analysis: do these sequence-sensitive measures provide insights

Received May 6, 1993; revised Dec. 15, 1993; accepted Jan. 27, 1994.

P.E.R. acknowledges support from NIH Grant NS19176. A.M.A. acknowledges partial support from the American Society for Engineering Education–Navy Sabbatical Support program. M.A.J.-M. acknowledges technical assistance of Oscar Chavoya-Aceves and Maricela Quintana-Lopez. We thank Tanya Schmah and the referees for several valuable comments.

Correspondence should be addressed to P. E. Rapp, Department of Physiology, The Medical College of Pennsylvania, 3300 Henry Avenue, Philadelphia, PA 19129.

Copyright © 1994 Society for Neuroscience 0270-6474/94/144731-09\$05.00/0

into categories of neural behavior that could not be obtained by a systematic statistical analysis of the interspike interval distribution? The results presented here suggest that the answer is yes. Analyses based on calculations of algorithmic complexity show that it is possible to identify distinct classes of neural behavior. Some spike trains display significant levels of sequence-sensitive structure while others are profoundly disordered. While this separation is readily apparent when complexity is measured, further calculations showed that the classification would not have been discovered if the examination was limited to the analysis of the interspike interval probability density function.

The usefulness of complexity measurements in neuroscience is not limited to the analysis of interspike intervals. Other measures of complexity have been applied to the analysis of the electroencephalogram (Wu and Xu, 1991; Xu and Wu, 1992) and to animal behavior (Paulus et al., 1990). Selz and Mandell (1993) have used measures of complexity to characterize experiments in which human subjects eliminate lattice points from a computer screen with a mouse-driven cursor. Pilot studies comparing results obtained with normal controls against subjects presenting a variety of personality disorders showed a correlation between complexity measures of behavior and diagnostic category.

Context-free grammar complexity

Several alternative definitions of complexity have been published (Kolmogorov, 1965; Zvonkin and Levin, 1970; Chaitin, 1974). The specific definition used here, the context-free grammar complexity, has been presented in greater detail by Jiménez-Montaño (Ebeling and Jiménez-Montaño, 1980; Jiménez-Montaño, 1984). This definition is an evolute of Kolmogorov's and Chaitin's, who held that the complexity of a message can be quantified by determining a measure of the length of the program required to generate the message. The procedure is best described by considering a specific example in which an upper bound of the complexity of a sequence of symbols is estimated. The choice of a procedure for reducing an experimentally obtained spike train to a sequence of symbols will then be considered. Alternative definitions of complexity and their possible advantages will be considered in the last section of this article.

Consider the following binary symbol sequence:

$$M = 101101011010001001$$

The object is to reduce this message to an instruction sequence constructed with the smallest possible number of symbols. We begin by searching for repeated pairs. the pair 0,1 is repeated six times in the message. The symbol a, a = 0,1, is introduced. The message becomes

$$M = 1 a 1 a 1 a 1 a 0 0 a 0 a.$$

The sequence 1,a is repeated three times. A new symbol is introduced:

$$a = 0 1$$

$$b = 1 a$$

$$M = b b a b 0 0 a 0 a$$

At this stage, repeated pairs have been exhausted. A search for repeated triples begins. None are found. The reduced message is

$$M = b^2 a b 0^2 a 0 a$$

$$a = 0 1$$

$$b = 1 a,$$

where exponentials are used to compress repeated symbols.

It is necessary to assign a quantitative measure to the amount of information encoded in a message that consists of symbols and their exponents. Since a limited number of symbols are used, each symbol can be encoded in a fixed number of bits. The exponents represent a special case since, in general, there is no upper bound on their potential size. The number of bits required to encode any integer as a binary string is, by definition, the smallest integer that is at least as large as the number's base 2 logarithm. A qualitative elaboration of this argument is given in Chaitin (1975). A more technical presentation is given in Chaitin (1974).

In the implementation used here, an upper bound of the complexity is determined by applying the following rules: each symbol in the sequence contributes 1 to the complexity; exponentials contribute logarithmically. In the final reduction seven symbols appear in *M* and the exponent 2 appears twice. *M* contributes $7 + 2 \log_2 2$ to the total. Symbols a and b each contribute 2. The complexity estimate is

$$\text{complexity}(M) = [7 + 2 \log_2 2 + 2 + 2] = 13.$$

The square brackets indicate that the integer part is to be taken.

The procedure generalizes immediately to symbol sequences composed of nonbinary alphabets. Consider

$$M = 1 1 2 3 1 1 4 2 3 1 1 4 4 2 3 3.$$

As before, the process begins with a search for repeated pairs. In this case the pair 1,1 and the pair 2,3 are repeated three times:

$$a = 1 1$$

$$b = 2 3$$

$$M = a b a 4 b a 4 4 b 3$$

The pair b,a is repeated, but it is only repeated twice. Calculations with simple examples can easily show that replacing a pair that is only repeated twice with a new symbol will not reduce the complexity. The procedure then searches for repeated triples. Triple b,a,4 is repeated twice. In the case of repeated triples, substitution does result in a reduction of complexity even if the triple only appears twice. The message becomes

$$M = a c^2 4 b 3$$

$$a = 1^2$$

$$b = 2 3$$

$$c = b a 4.$$

There are no other repeated triples in *M*. There are no repeated four-symbol sequences. The reduction has converged. The contributions of *M*, a, b, and c to the estimate are $5 + \log_2 2$, 1 + $\log_2 2$, 2, and 3, respectively:

$$\text{complexity}(M) = [5 + \log_2 2 + 1 + \log_2 2 + 2 + 3] = 13$$

We have recently learned that many of the ideas incorporated into this compression algorithm were anticipated in a series of reports by Wolff (1975, 1976, 1977, 1987). Wolff explicitly addresses psychological issues and has argued that data compression following this procedure can successfully model speech segmentation, concept formation, and language acquisition. Our

application of the algorithm is limited to its use as an empirical metric of complexity.

A qualitative understanding of this estimate of complexity can be obtained by considering artificial examples. For a fixed number of symbols, the lowest complexity would be obtained with a message consisting of a single repeated symbol. An estimate of 22 is obtained if the same symbol is repeated 1000 times. The next simplest cases would be periodic messages constructed with repeated symbol sequences. One thousand element data sets constructed by repeating 2, 4, and 8 element symbol strings have complexities of 22, 24, and 30, respectively. At the other extreme, the highest complexity estimates will be obtained with messages constructed using random numbers. The results presented in Table 1 serve two purposes. First, they determine the upper end of the range of binary complexity obtainable with 1000 elements. Second, these artificial examples provide a way of addressing the question of how data should be reduced to a sequence of symbols. We examined artificially generated random numbers with five different distributions. The random number generator follows a design of Knuth's (1981) as implemented by Press et al. (1986). As a further check, we repeated the calculations using a natural source of random numbers. The cobalt data were obtained by measuring the time intervals between decays in a sample of ^{60}Co .

Before the complexity can be estimated using our algorithm, it is necessary to reduce these data to a sequence of symbols. The usefulness of subsequent complexity calculations depends crucially on the procedure used to partition the data among a finite alphabet of symbols. If this is done inappropriately, spurious results will be obtained. The computational results displayed in Table 1 were obtained after the data were reduced to a binary sequence by partitioning about the mean, the median, and the midpoint. If the data value was less than the mean, it was assigned the symbol 0. If it was greater than the mean, symbol 1 was assigned. Partitions about the median and the midpoint followed the same procedure. Five independent data sets were used from each distribution. The average complexity is displayed with its SD. It is seen that partitioning asymmetrically distributed data about the midpoint can give a misleading indication of structure. The best partition can be operationally defined as the partition that most effectively reveals the randomness of the original data. That is, the best partition is the one that gives the largest complexity estimate. These empirical results indicate that partitioning about the median meets this criterion. All binary symbol sequences reported in this article were constructed by partitioning about the median.

It is possible to use calculations like those in Table 1 to estimate the uncertainty in complexity estimates. Ten data sets were drawn from each of three stationary processes: (1) artificially generated, uniformly distributed random numbers; (2) data generated by a three-dimensional generalization of the Hénon difference equations; and (3) the cobalt data. (The same equipment is used for cobalt measurements and for the interspike interval measurements.) The average complexity is determined for each group of 10 data sets. The SD divided by the average complexity, expressed as a percentage, provides a rough estimate of the uncertainty in the determination of complexity. Uncertainty is found to vary with the data source and with the size of the data set. Calculations were performed with 50, 100, 200, and 1000 element data sets. The corresponding uncertainties obtained with random numbers were 5.7%, 2.9%, 2.0%, and 1.2%. Using the cobalt data, the uncertainties were 5.1%, 4.5%,

Table 1. Average complexity of 1000 point data sets ($N = 5$)

Distribution	Symbol sequence formed about the mean	Symbol sequence formed about the median	Symbol sequence formed about the midpoint
Uniform	273 ± 2.1	272 ± 2.9	276 ± 2.0
Gaussian	274 ± 5.0	274 ± 4.7	271 ± 3.0
Exponential	263 ± 4.0	272 ± 2.9	57 ± 16.0
Bimodal	274 ± 5.7	274 ± 5.4	274 ± 5.4
Poisson	275 ± 2.4	276 ± 4.6	271 ± 4.8
Cobalt	267 ± 3.6	275 ± 2.5	94 ± 13.6

Calculations with messages obtained by reducing random numbers to symbol sequences establish the upper bound of complexity that can be obtained with 1000 elements. Binary sequences produced by partitioning about the median are seen to be insensitive to the distribution. The results are reported with SDs obtained from calculations using five data sets for each distribution.

2.6%, and 0.9%, and with the Hénon data they were 9.9%, 7.6%, 6.7%, and 1.7%. Most of the calculations in this study were performed with 1000 event data sets where the estimated uncertainty is on the order of 2%. The smallest data sets used were records obtained in 60 sec epochs. On average there were 503 events in these data sets. The smallest has 106 elements.

Experimental Methods and Results

Single-unit records were obtained from cortical neurons of the rat before and after the application of penicillin to the cortex in a procedure developed by Matsumoto and Ajmone-Marsan (1964) and Prince (1968). Adult, male, albino rats (Sprague-Dawley strain) in the weight range of 200–350 gm were anesthetized with chloral hydrate. Animals were restrained by means of a stereotaxic headholder. The skin and cutaneous muscles were incised on the dorsal midline from the nasion to the occiput, and the temporal muscles and the periosteum were reflected from the underlying cranium. Initial measurements from the skull using the bregma as the origin (Zeman and Innes, 1963) served as an aid to approximate the location of the underlying somatosensory and motor cortex. A small portion of bone overlying the cortex was removed by a drill until only a thin layer of bone remained. This thin layer was removed by chipping away the bone with a fine forceps along the circumference of the cranial defect. Saline was applied to the cortex throughout the procedure. Following surgery, animals were maintained on anesthesia as determined by monitoring the electrocorticogram, corneal and withdrawal reflexes, and pupillary dilation. Supplementary doses were administered via an intravenous canula.

Recordings were performed in a Faraday cage. Single-unit records were obtained with insulated tungsten microelectrodes with a resistance of approximately 10 M Ω . The electrode was positioned over the cortex in a micromanipulator and lowered under visual guidance to avoid damage to the middle cerebral artery and its branches. The frequency response of the amplifier was adjusted to each recording situation. Typically, we set the low-frequency filter at 60 Hz and the high-frequency setting at 60 kHz. The micromanipulators, stereotaxis equipment, preamp, amplifiers, oscilloscope, and cage were connected to a common building ground. Interspike intervals were measured to an accuracy of 10 μsec . In Matsumoto's implementation of the procedure the epileptogenic focus is created by placing 1 m 2 of filter paper saturated with aqueous penicillin (100,000–200,000 U/ml penicillin G) directly on the surgically exposed cerebral cortex. We want explicitly to note that the penicillin focus is an im-

Table 2. Statistical properties of order-sensitive measures, 1000 events in each data set

Neuron	C_2 Spon.	C_2 Pen.	RC_2 Spon.	RC_2 Pen.
1	227	266	1.622	2.454
2	276	274	0.320	0.454
3	266	270	2.952	3.667
4	275	275	1.991	3.614
5	227	269	0.899	1.671
6	223	243	2.367	3.160
7	258	274	1.167	3.506
Average	252	267	1.617	2.647
SD	22.2	11.2	0.902	1.209

Significance testing (spontaneous vs penicillin treated)

C_2 : $t = 2.277$, $p = 0.063$

RC_2 : $t = 3.773$, $p = 0.009$

The values of binary complexity, C_2 , and the rate of complexity production, RC_2 , obtained from 1000 element spike trains partitioned about the median are displayed. A paired t test was used to compare the values obtained in the spontaneous and the penicillin-treated case.

perfect model of human focal epilepsy (Schwartz et al., 1970; Calvin et al., 1973; Wyler et al., 1982). The method does, however, provide a simple method for producing at least an approximation of a focal seizure in an acute preparation.

All neurons showed an increased firing rate in response to penicillin. The SD and the average deviation decreased in the penicillin condition. These changes were significant (mean, $p = 0.035$; average deviation, $p = 0.018$; SD, $p = 0.012$, where the null hypothesis is that the differences observed between penicillin-induced behavior and spontaneous behavior are due to random variation). The statistical separation between spontaneous and penicillin-treated behavior obtained by the mean and by the SD defines the benchmarks against which other measures, specifically complexity, should be compared. Other measures of the distribution, notably the third-order moment (skewness) and the fourth-order moment (kurtosis), do not show significant change. The distributions of interspike intervals display the long tails characteristic of bursting neurons. The examination of higher moments cannot be justified in these cases. Wyler et al. (1978) define the burst index of a spike train as the percentage of interspike intervals less than 5 msec. We calculated the index using 3, 4, and 5 msec as the defining criterion. No significant differences were observed between the spontaneous and penicillin-treated spike trains ($p = 0.241$, $p = 0.518$, and $p = 0.360$, respectively).

Binary complexity

Binary complexity, denoted by C_2 , is the value of complexity obtained when the interspike interval data are reduced to a sequence of zeros and ones about the median. The results obtained with the 1000 event data sets are reported in Table 2. It is seen that, on average, the complexity increases during a penicillin seizure. However, the increase does not occur in all neurons and it is not markedly significant ($p = 0.06$). It is important to remember that the results shown in the first two columns of the table are obtained with uniformly sized data sets (1000 events). Because the mean firing rate differs from neuron to neuron, these records cover markedly different lengths of time. Because timely responsiveness is an essential property of any successful biological system, it could be argued that complexity

itself is not an appropriate measure. Rather, complexity generation per unit time may be a better measure of biological complexity. RC_2 is defined as the rate at which complexity is generated. It is calculated by dividing the complexity by the corresponding time required by that neuron to fire the 1000 action potentials. When this is done the distinction between spontaneous and seizure behavior is significant. The probability of the null hypothesis drops to $p = 0.009$.

We performed calculations comparing complexities presented by fixed epoch data sets for 60, 120, and 180 sec. In each case the differences between conditions are significant, $p = 0.008$, $p = 0.001$, and $p = 0.004$, respectively. A comparison with the distribution-determined significance tests is of interest. In each case (60 sec, 120 sec, 180 sec) the probability of the null hypothesis obtained with fixed epoch complexity is smaller than the smallest value obtained from the distribution. In one case, 120 sec, the value of p is less than 1/10 of the smallest value obtained from the statistical examination of the distribution. Thus, the previously specified statistical benchmarks have been exceeded. Complexity is more effective in distinguishing between spontaneous and seizure behavior than classical measures of the distribution of the interspike interval histogram.

However, the results obtained by calculating RC_2 seem more impressive than they really are. It should be remembered that the mean interspike interval decreases in response to penicillin. RC_2 is obtained by dividing the value of complexity by its corresponding 1000 event epoch length. We must therefore ask if the impressive statistical separation seen in RC_2 occurs because penicillin-treated neurons fire faster. Two properties of the spike train contribute to the complexity: (1) the number of symbols in the message, which is determined by the firing rate, and (2) the structure of the message. The high correlation between fixed epoch complexity and the mean interspike interval (as determined by both Kendall's τ and the Spearman rank order coefficient) indicates that, for these neurons, the first mechanism is an important determinant of complexity. However, even in these cases, firing rate does not completely determine complexity. This is demonstrated definitively by the calculations with surrogate data presented in a subsequent section.

Complexity measured with expanded symbol alphabets

The definition of complexity used here is not limited to binary symbol sequences. It can be applied to sequences composed with larger symbol sets. The results in Table 1 argue for partitioning about the median when constructing binary symbol sequences. This procedure readily generalizes to expanded symbol sets. An equal number of interspike intervals will be assigned to each symbol. For example, if four symbols are used, the shortest 25% of the intervals will be assigned symbol 0. The next 25% are assigned symbol 1, and so on.

The complexity was calculated using 180 sec epoch data sets and 4, 6, 8, and 10 symbol alphabets. In each case complexity increased in response to the administration of penicillin. Significance levels were $p = 0.003$, $p = 0.003$, $p = 0.004$, and $p = 0.005$, respectively. Contrary to our anticipations, C_{10} is not a significantly better measure than C_2 . When 1000 event data sets were examined with expanded symbol sets, it was found that the ability of complexity measures to discriminate between spontaneous and penicillin-treated behavior actually deteriorated as the number of symbols increased. For 2, 4, 6, 8, and 10 symbol alphabets $p = 0.063$, 0.090, 0.135, 0.102, and 0.172, respectively.

With these data it is seen that C_{10} is less successful than C_2 in discriminating between the two physiological conditions. We believe that we now understand why this is the case. Consider the limiting value when N_s , the number of symbols in the alphabet, approaches N_{DATA} , the number of interspike intervals. In this limit each interspike interval is assigned to its own symbol. The resulting symbol sequence consists of N_{DATA} unique symbols and has a complexity equal to N_{DATA} . Thus, as N_s increases, complexity simply becomes a measure of the size of the data set. Increasing N_s is seen to be subject to a law of diminishing returns. The question is how quickly does the limit become important; how rapidly does the value of complexity approach N_{DATA} as N_s increases? This question was investigated by examining the value of complexity obtained from sets of random numbers where in each case $N_{\text{DATA}} = 1000$. Five sets of uniform deviates were investigated. The average value of complexity should approach 1000 as N_s increases. When N_s is equal to 10, complexity is equal to 699; that is, complexity has reached 70% of its limiting value when N_s is 1% of N_{DATA} .

The preceding argument indicates why increasing N_s will not necessarily be helpful. However, the results make an even stronger point. Increasing N_s can actually result in a deterioration in discrimination. The reason for this has to do with the size of the data sets. Experimental calculations suggest that much larger data sets are required for meaningful complexity calculations if N_s is greater than 2. For this reason all subsequent calculations in this article will focus on the $N_s = 2$ case.

Calculations with expanded symbol alphabets point out another limitation of this definition of complexity. Consider the case where $N_s = 10$. Each symbol in the alphabet is treated as an arbitrary symbol. Symbol 5 does not "know" that it is between symbols 4 and 6. Similarly, it does not "know" that it is very far from symbols 1 and 10. All of this structure is lost to this definition of complexity.

Comparison with algorithm zero surrogates

In the discussion of the binary complexity of fixed epoch data sets, $C_2(60 \text{ sec})$, $C_2(120 \text{ sec})$, $C_2(180 \text{ sec})$, we identified two properties of the spike trains that contribute to the complexity: the number of symbols in the message, which is determined by firing rate; and the structure of the message. We now address the following question: is the pattern of the message a significant contributor to C_2 , or is C_2 simply an exotic way of measuring the firing rate? If the complexity is completely determined by frequency, there is little reason to compute it.

This question can be addressed with a very simple calculation. Consider as a specific example the 1000 event spike train that was obtained from neuron 1 prior to the application of penicillin. This data set was found to have a binary complexity of $C_2 = 227$. The spike train is defined by the sequence $I_1, I_2, I_3, \dots, I_{1000}$. This sequence was subjected to a random shuffle, and the complexity of the shuffled sequence was measured. It was found to be $C_2 = 271$. This is a value typical of the results obtained with random numbers reported in Table 1. The original data set and the randomly shuffled data set have the same distribution. The average firing rate is identical for each data set. Yet, the complexity values are different. This indicates that the lower value of complexity obtained with the original data is due to the internal structure of the message, a structure that was destroyed by the random shuffle. Is the difference between the value of complexity obtained with the original data and the shuffled variant significant? This question can be addressed by

performing this test with several different random shuffles of the original data.

The method of surrogate data defines procedures for performing these tests in a systematic way (see Theiler et al., 1992, for a review; surrogate data methods have also been used by Longtin, 1993b, in his study of periodically forced sensory neurons). The comparison of measures obtained with original data and with a randomly shuffled variant of the original data is called the Algorithm Zero surrogate test. Given the original data set, a measurement is performed giving the result M_{orig} . In our case we are measuring complexity, but this reasoning can be applied to any measure. Using different random shuffles, several shuffled data sets are produced. These are called Algorithm Zero surrogate data sets. The same measurement is performed with each surrogate. $\langle M_{\text{surr}} \rangle$ denotes the average measurement obtained with surrogate data. We then ask, is M_{orig} significantly different from $\langle M_{\text{surr}} \rangle$? When addressing this question we want a measure that incorporates a sensitivity to possible variability in measurement of M . Let σ_{surr} denote the SD of the measurements obtained with surrogates. Following Theiler et al. (1992), S is defined by

$$S = |M_{\text{orig}} - \langle M_{\text{surr}} \rangle| / \sigma_{\text{surr}}$$

S gives the number of SDs separating the value of the measurement obtained with the original data and its surrogates.

The randomly shuffled surrogate examines a specific null hypothesis: there are no temporal correlations in the spike train; that is, the structure of the spike train is indistinguishable from uncorrelated noise that has the same distribution. If the complexity was completely specified by the frequency, then S would be identically zero. Table 3 shows the value of S obtained when the binary complexity is measured ($M = C_2$) for fixed epoch spike trains of 60, 120, and 180 sec and for 1000 event spike trains. Twenty surrogates were used in each calculation.

On examining the results in Table 3 several important observations can be made. First, S is not always equal to, or near, zero. For example, the average value of S for 1000 event records obtained prior to the administration of penicillin is 9.4. Recalling the definition of S (the number of SDs from the expectation value of the null hypothesis), it is seen that spontaneous neural spike trains possess a very high degree of structure that is determined independently of frequency. Second, it is seen that in most cases S decreases in response to penicillin. This is equivalent to stating that the behavior becomes less distinguishable from uncorrelated noise during penicillin seizures. Third, by comparing 60 sec, 120 sec, and 180 sec calculations it is seen that, as expected, the value of S varies with time. This suggests that complexity in combination with surrogate data calculations could be used to monitor time-dependent changes in neural behavior. This possibility will be explored in a subsequent publication.

An examination of the S values in Table 3 indicates that some spike trains are essentially indistinguishable from their randomly shuffled surrogates, while some have a significant degree of internal structure. For example, consider the 1000 event spike trains obtained after the administration of penicillin. The S values fall into two very distinct classes. When viewed in 1000 event segments, the spike trains from neurons 2, 4, and 7 are disordered, while spike trains from neurons 1, 3, 5, and 6 show some structure. We now ask a question that is crucial to an assessment of the value of this analysis: could the classification of neural behavior based on the results in Table 3 have been

Table 3. Comparisons with Algorithm Zero surrogate data

Neuron	60 sec	120 sec	180 sec	1000 events
<i>S</i> values: spontaneous behavior				
1	10.0	13.0	14.2	21.3
2	1.1	1.0	1.0	0.2
3	0.7	2.0	1.9	2.8
4	0.7	0.1	0.0	0.1
5	5.6	7.0	8.4	23.3
6	9.7	11.5	10.4	12.4
7	2.5	5.7	3.3	5.4
Ave.	4.3	5.8	5.6	9.4
SD	4.1	5.1	5.4	9.8
<i>S</i> values: penicillin treated				
1	0.5	0.1	2.1	2.7
2	0.3	1.0	1.7	0.6
3	5.5	1.1	1.8	2.2
4	0.0	0.8	0.3	0.4
5	3.4	0.2	0.6	1.7
6	8.3	11.0	16.0	9.9
7	0.9	0.4	1.4	0.1
Ave.	2.7	2.1	3.4	2.5
SD	3.2	3.9	5.6	3.4

The binary complexity of an Algorithm Zero surrogate is determined by randomly shuffling the original data, partitioning the shuffled data set about the median, and calculating the complexity of the resulting symbol sequence. The *S* value is a measure of the difference between the complexity of the original data set and its surrogates. A high value of *S* indicates a high degree of structure, as assessed by this class of surrogate, in the original data.

made by examining the distributions of the 1000 event data sets? This question is addressed in Table 4. The comparison of the average values obtained from each group and the separation ratio (the larger value divided by the smaller) is displayed in Table 4. The average value of *S* from each group differs by a factor of 14. None of the measures obtained from the statistical properties of the distribution approaches this separation. Analogous results were obtained by Selz and Mandell (1991), who found that a measure of complexity similar to the one described here could discriminate three classes of intermittently firing brainstem neurons more clearly than the higher moments of the interspike interval distribution.

Because randomly shuffled surrogates have the same distributions as the original data, the classification of neural behavior based on complexity would not have been made if the analysis had been limited to distribution-determined measures. We can now return to a question raised in the introductory remarks: do sequence-sensitive measures provide insights into categories of neural behavior that could not be obtained by a systematic statistical analysis of the distribution? The answer appears to be yes.

Comparisons with Algorithm One and Algorithm Two surrogates

The analysis with surrogate data is based on the underlying assumption that the frequency of the membrane potential, and hence the spike train, is determined by a dynamical system controlled by the membrane's chemical environment. Algorithm Zero surrogates determine if the function specifying the frequency of the membrane potential is indistinguishable from

Table 4. Comparison of the statistical properties of structured and disordered spike trains

	Structured neurons (average values)	Disordered neurons (average values)	Separation ratio
<i>S</i> values (Algorithm 0)	3.80 ± 2.81	0.27 ± 0.15	14.1
Mean	104.99 ± 40.51	252.73 ± 304.16	2.4
Median	22.18 ± 12.93	33.17 ± 19.01	1.5
Average deviation	119.77 ± 51.73	300.69 ± 400.53	2.5
Standard deviation	182.59 ± 73.54	403.61 ± 522.60	2.2
Skewness	3.05 ± 0.51	2.47 ± 0.53	1.2
Kurtosis	11.83 ± 3.51	8.25 ± 3.94	1.4
Coefficient of variation	174.79 ± 20.61	143.59 ± 24.82	1.2

The *S* values for the 1000 event, penicillin-treated spike trains indicates that the records obtained from neurons 2, 4, and 7 are disordered, while those of neurons 1, 3, 5, and 6 display some structure. The average *S* value of each group is separated by a factor of 14.1. Statistical measures of the distributions differ by, at most, a factor of 2.5. The classification of neurons (disordered/structured) made on the basis of complexity calculations combined with surrogate data calculations would not, therefore, have been made on the basis of an examination of the corresponding interspike interval distributions.

noise. This class of surrogates detects the presence of any structure in the data. Suppose random numbers were filtered by a linear filter. A smoothly varying waveform results. (An example constructed from filtered noise and its analysis with surrogates is given in Rapp et al., 1993a.) If a random shuffle is applied to the filter's output, this smooth structure is destroyed. Applied to this signal, Algorithm Zero surrogates give a large value of *S* even though the underlying dynamical behavior is random. Algorithm One surrogates explicitly address this possibility. Several investigators independently presented an elegant algorithm for investigating the following null hypothesis: the signal was produced by the action of a linear filter on noise. A review of the previous literature is given in Theiler et al. (1992). This is a very pertinent concern in a neurophysiological application because the membrane acts as a filter. The time constants of cortical neurons are on the order of 10–20 msec (Stratford et al., 1989; Douglas and Martin, 1990). The neural membrane effectively acts as a low-pass filter and severely attenuates signals with a frequency in excess of 15 Hz.

The construction of Algorithm One surrogates is based on the following arguments. A linear filter is specified by its transfer function, and the spectrum of linearly filtered noise will be the spectrum of this transfer function. The Algorithm One null hypothesis is therefore equivalent to stating that all of the signal's structure is specified by its spectrum. This hypothesis can be tested by constructing surrogates that are generated from random numbers, but which have a spectrum identical to the original data. Algorithm One surrogates are produced in a three step process. (1) The Fourier transform of the original data set is determined. (2) The phases of the Fourier transform are randomized. (3) A surrogate data set is constructed by taking the inverse transform. Since the phase does not contribute to the spectrum, the surrogates and the original data set have identical spectra.

The efficacy of Algorithm One surrogates is limited to the action of linear filters. Consider the following hypothetical case. Suppose time series $\{y_j\}$ was produced by the action of a linear filter on a set of random numbers, and that $\{z_j\}$ is generated by the rule $z_j = h(y_j)$, where h is a static nonlinear function. Because

h is nonlinear, the value of S obtained from Algorithm One surrogates of $\{z_j\}$ will be significantly greater than zero, even though the underlying dynamical structure is random. Algorithm Two surrogates, introduced by Theiler et al. (1992), address the following null hypothesis: the time series is linearly correlated noise that has been transformed by a static, monotonic nonlinear function.

The first step in constructing Algorithm Two surrogates is to approximate h^{-1} ; $y = h^{-1}(x)$, where $\{x_j\}$ is the original time series. The elements of $\{y_j\}$ are drawn from a random Gaussian distribution that has the same rank ordering as $\{x_j\}$. (Two time series have the same rank ordering if the j th element is the k th largest in both series.) The previous algorithm can be applied to $\{y_j\}$ to produce $\{y'_j\}$, an Algorithm One surrogate of $\{y_j\}$. Time series $\{x'_j\}$ is obtained by shuffling $\{x_j\}$ so that it has the rank structure of $\{y'_j\}$. By drawing element of $\{x'_j\}$ from set $\{x_j\}$, but retaining the rank structure of $\{y'_j\}$, we have approximated the effect of h acting on $\{y'_j\}$; $x'_j = h(y'_j)$. Time series $\{x'_j\}$ is therefore an Algorithm Two surrogate of $\{x_j\}$.

The distinction between the null hypotheses of Algorithm One and Algorithm Two may seem insubstantial. It is, however, of considerable practical significance. Unlike Algorithm One, Algorithm Two preserves the distribution of the original data. Using this property, it is possible to construct examples of appropriately transformed random systems where Algorithm One gives very high values of S and Algorithm Two returns a value of S effectively equal to zero. We have constructed an example using random numbers and the complexity measure used in this article in which Algorithm One gives $S = 31.8$ and Algorithm Two gives $S = 0.6$ (Rapp et al., 1993b). Similar examples can be produced by signals transformed by the logarithmic amplifiers commonly used in many experimental systems. Examinations limited to Algorithm One can give false positive indications of meaningful structure where none in fact exists. The distinction between linear and nonlinear transformations is relevant to a neurophysiological application since a membrane's response to effectors is nonlinear. In the computational example investigated in Rapp et al. (1993a), the function h is the allosteric ligand binding function. Randomly constructed time series transformed by this function give significant values of S with Algorithm One and negligible values with Algorithm Two.

Results obtained with the 1000 event data sets are presented in Table 5. S_{A0} denotes the value of S obtained with Algorithm Zero surrogates. S_{A1} and S_{A2} are the values obtained with Algorithm One and Algorithm Two. Twenty surrogates were used in each of these calculations. The emerging typical pattern is $S_{A0} > S_{A1} > S_{A2}$. As anticipated, Algorithm 2 is a much more stringent test of structure. However, even with Algorithm Two, the calculations confirm that at least some neural spike trains have significant nonlinear dynamical structure. The calculations also indicate that the behavior becomes significantly more disordered during penicillin seizures.

Discussion

We begin the discussion with a summary of the principal results. The context-free grammar complexity of single-unit cortical spike trains was found to increase during penicillin-induced focal seizures. Greater statistical discrimination between the two physiological conditions (spontaneous behavior vs penicillin-altered behavior) was obtained with the complexity than by measures such as the mean interspike interval, the SD, skewness, and kurtosis. An examination of fixed epoch data sets showed that

Table 5. Comparisons with surrogate data, 1000 events in each data set

Neuron	S_{A0} Spon.	S_{A1} Spon.	S_{A2} Spon.	S_{A0} Pen.	S_{A1} Pen.	S_{A2} Pen.
1	21.3	13.4	9.0	2.7	2.6	1.8
2	0.2	0.5	0.2	0.6	0.2	0.3
3	2.8	2.2	1.9	2.2	0.7	0.2
4	0.1	0.5	0.1	0.4	0.3	0.1
5	23.3	12.9	8.0	1.7	1.3	1.4
6	12.4	11.5	6.0	9.9	9.5	4.3
7	5.4	5.4	3.5	0.1	0.4	0.4
Ave.	9.4	6.6	4.1	2.5	2.1	1.2
SD	9.8	5.8	3.6	3.4	3.4	1.5

Significance testing (spontaneous vs penicillin treated)

$$S_{A0}: t = 1.945, p = 0.100$$

$$S_{A1}: t = 2.442, p = 0.050$$

$$S_{A2}: t = 2.581, p = 0.042$$

S_{A0} denotes the S value obtained with Algorithm Zero (random shuffle) surrogates. S_{A1} denotes the value obtained with Algorithm One (random phase) surrogates, and S_{A2} denotes the value obtained with Algorithm Two (Gaussian scaled) surrogates. The sequence of algorithms applies increasingly demanding tests of structure to the spike trains. The significance test (paired t test) indicates that though the S values obtained with Algorithm Two are low, the values obtained in the spontaneous records are significantly greater than those obtained after the application of penicillin ($p = 0.042$).

two factors contribute to the complexity: the firing rate and the internal structure of the spike train. However, calculations with randomly shuffled surrogates of the original data showed that the complexity is not completely determined by firing rate. The sequence-sensitive structure of the spike train is a significant contributor.

Spontaneous neural spike trains were found, on average, to possess a high degree of structure that was lost by random shuffles of the original data. This structure typically decreased in response to penicillin. For any given neuron, the degree of order in its output is not fixed, but rather varies through time. By combining complexity measurements with surrogate data calculations, it is possible to classify neurons according to the dynamical structure into distinct groups of ordered or disordered spike trains. This classification could not have been made on the basis of distribution-determined measures. Computations with more sophisticated kinds of surrogate data show that the structure observed using complexity measures cannot be attributed to linearly correlated noise or to linearly correlated noise transformed by a static, monotonic nonlinearity. The patterns in spike trains appear to reflect genuine nonlinear structure.

Alternative mathematical definitions of complexity should be investigated. The definition used here gives very high values for random structures such as an ideal gas, and very low values for highly ordered structures like a crystal. It has been argued (Huberman and Hogg, 1986) that a better measure of complexity would give the highest score for structures intermediate to random gases and perfect crystals. These issues are addressed by the measure of complexity introduced by Huberman and Hogg (1986). They consider the complexity of tree structures. This is, however, applicable to our data since every binary sequence of zeros and ones can be restated as a network. Similar ideas have been implemented in computational ϵ -machines (Crutchfield and Young, 1989). The complexity measure introduced by Crutchfield and Young gives low values for periodic and for

random systems and higher values for chaotic systems. This is of particular relevance to these data since there have been a number of previous reports of nonlinear structure in neural spike train data (Rapp et al., 1985; Mpitsos et al., 1988; Longtin et al., 1993b). However, when considering measures of biological behavior it should be remembered that, somewhat paradoxically, the most sensitive measures of dynamical behavior are not necessarily the best. Sophisticated dynamical measures, for example, correlation dimension, are extremely sensitive to noise in the data. They can only be reliably applied when large, noise-free data sets are available (Eckmann and Ruelle, 1992; Rapp, 1992). A more robust measure like the grammar complexity may prove more successful in discriminating between different physiological states.

Independently of technical considerations about the definition of complexity, there may be fundamental biological limitations to what can be learned from the analysis of the structure of interspike interval spike trains. A summary of the arguments leading to this conclusion has been presented by Douglas and Martin (1991). A neuron receives inputs from hundreds of neurons. The mean interspike interval is on the order of 100 msec ($\omega = 10$ Hz). The fine structure of a signal from any given input neuron would be lost if it were encoded against this low-frequency background. These considerations would argue for encoding temporally structured signals at higher frequencies. However, as previously noted, the neural membrane effectively acts as a low-pass filter. In the absence of an alternative coding mechanism, for example, stochastic resonance (Chialvo and Apkarian, 1993; Longtin, 1993a), these arguments suggest that it would be difficult to construct a fine structure coding scheme that could be interpreted by the recipient neuron. These considerations place an important constraint on the implications of the results presented in this article. If limiting arguments of this type are valid, it would be very difficult to construct a theory of neural computation based on the fine structure of single neuron spike trains. While bearing this possible limit in mind, we should not lose sight of two points. First, some structure is often present. This is demonstrated by large S values obtained in comparisons with surrogate data. Second, the internal structure of the spike trains changed significantly with a change in physiological state. Therefore, if a specific goal is to characterize quantitatively changes in neural behavior associated with pathological states, these measures have potential value as metrics of behaviors that cannot be detected by an examination of the interspike interval distribution.

References

- Abeles M, Gerstein GL (1988) Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J Neurophysiol* 60:909–924.
- Calvin WH, Ojemann GA, Ward AA (1973) Human cortical neurons in epileptogenic foci: comparison of interictal firing patterns to those of “epileptic” neurons in animals. *Electroencephalogr Clin Neurophysiol* 34:337–351.
- Chaitin GJ (1974) Information theoretical computational complexity. *IEEE Trans Inform Theory* IT20:10–15.
- Chaitin GJ (1975) Randomness and mathematical proof. *Sci Am* 232(5):47–52.
- Chen Y-Q, Ku Y-H (1992) Properties of favored patterns in spontaneous spike trains and responses of favored patterns to electroacupuncture in evoked trains. *Brain Res* 578:297–304.
- Chialvo DR, Apkarian AR (1993) Modulated noisy biological dynamics: three examples. *J Stat Phys* 70:375–391.
- Crutchfield JP, Packard NH (1983) Symbolic dynamics of noisy chaos. *Physica* 7D:201–223.
- Crutchfield JP, Young K (1989) Inferring statistical complexity. *Phys Rev Lett* 63:105–108.
- Dayhoff JE (1984) Distinguished words in data sequences: analysis and applications to neural coding and other fields. *Bull Math Biol* 46:529–543.
- Dayhoff JE, Gerstein GL (1983a) Favored patterns in spike trains. I. Detection. *J Neurophysiol* 49:1334–1346.
- Dayhoff JE, Gerstein GL (1983b) Favored patterns in spike trains. II. Application. *J Neurophysiol* 49:1347–1363.
- Douglas RJ, Martin KAC (1990) Neocortex. In: *The synaptic organization of the brain* (Sheppard GM, ed), pp 389–438. New York: Oxford UP.
- Douglas RJ, Martin KAC (1991) Opening the grey box. *Trends Neurosci* 14:286–293.
- Ebeling W, Jiménez-Montaña MA (1980) On grammars, complexity and information measures of biological macromolecules. *Math Biosci* 52:53–71.
- Eckmann J-P, Ruelle D (1992) Fundamental limitations for estimating dimensions and Liapunov exponents in dynamical systems. *Physica* 56D:185–187.
- Gerstein GL, Perkel DH, Dayhoff JE (1985) Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement. *J Neurosci* 5:881–889.
- Huberman BA, Hogg T (1986) Complexity and adaptation. *Physica* 22D:376–384.
- Jiménez-Montano MA (1984) On the syntactic structure of protein sequences and the concept of complexity. *Bull Math Biol* 46:641–659.
- Klemm KR, Sherry CJ (1982) Do neurons process information by relative intervals in spike trains? *Neurosci Biobehav Rev* 6:429–437.
- Knuth DE (1981) *The art of computer programming*, Vol 2, *Seminumerical algorithms*. Reading, MA: Addison Wesley.
- Kolmogorov AN (1958) A metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl Acad Nauk USSR* 119:861–864 [English summary: *Math Rev* 21:386].
- Kolmogorov AN (1965) Three approaches to the definition of the concept of quantity of information. *IEEE Trans Inform Theory* IT14:662–669.
- Legendy CR, Salzman M (1985) Bursts and recurrence of bursts in the spike trains of spontaneously active cortex neurons. *J Neurophysiol* 53:926–939.
- Lestienne R, Strehler BL (1987) Time structure and stimulus dependence of precisely replicating patterns present in monkey cortical neuronal spike-trains. *Brain Res* 437:214–238.
- Longtin A (1993a) Stochastic resonance in neuron models. *J Stat Phys* 70:309–326.
- Longtin A (1993b) Nonlinear forecasting of spike trains from sensory neurons. *Int J Bifurcation Chaos* 3:651–662.
- Mandell AJ, Selz KA (1993a) Brain stem neuronal noise and neocortical resonance. *J Stat Phys* 70:355–373.
- Mandell AJ, Selz KA (1993b) Resonance, synchronization and lexical redundancy in the expanding dynamics of brain stem neurons. *SPIE* 2036: in press.
- Matsumoto H, Ajmone-Marsan C (1964) Cortical cellular phenomena in experimental epilepsy. *Exp Neurol* 9:286–304.
- Mpitsos GJ, Burton RM, Creech HC, Soinilla SO (1988) Evidence for chaos in spike trains from neurons that generate rhythmic motor patterns. *Brain Res Bull* 21:529–538.
- Paulus MP, Geyer MA, Gold LH, Mandell AJ (1990) Application of entropy measures derived from the ergodic theory of dynamical systems to rat locomotor behavior. *Proc Natl Acad Sci USA* 87:723–727.
- Perkel DH, Bullock TH (1968) Neural coding. *Neurosci Res Prog Bull* 6(3).
- Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) *Numerical recipes. The art of scientific computing*. Cambridge: Cambridge UP.
- Prince DA (1968) The depolarization shift in epileptic neurons. *Exp Neurol* 21:467–485.
- Rapp PE (1992) Chaos in the neurosciences: cautionary tales from the frontier. *Biologist* 40:89–94.
- Rapp PE, Zimmerman ID, Albano AM, deGuzman GC, Greenbaum NN (1985) Dynamics of spontaneous neural activity in the simian motor cortex: the dimension of chaotic neurons. *Phys Lett* 110A:335–338.

- Rapp PE, Albano AM, Schmah TI, Farwell LA (1993a) Filtered noise can mimic low dimensional chaotic attractors. *Phys Rev* 47E:2289–2297.
- Rapp PE, Albano AM, Zimmerman ID, Jiménez-Montaño MA (1993b) Phase-randomized surrogates can produce spurious identifications of non-random structure. *Phys Lett*, submitted.
- Schmidt EM, Matsuga N, MacIntosh JS (1976) Chronic recording of neurons in epileptogenic foci of monkeys during seizures. *Exp Neurol* 52:459–466.
- Schwartz OR, Broggi G, Pappas GD (1970) Fine structure of cat hippocampus during sustained seizures. *Brain Res* 18:176–180.
- Selz KA, Mandell AJ (1991) Bernoulli partition-equivalence of intermittent neuronal discharge patterns. *Int J Bifurcation Chaos* 1:717–722.
- Selz KA, Mandell AJ (1992) Critical coherence and characteristic times in brain stem neuronal discharge patterns. In: *Single neuron computation* (McKenna T, Davis J, Zornetzer SF, eds), pp 525–560. New York: Academic.
- Selz KA, Mandell AJ (1993) Hierarchical lexical information estimates: when measure theoretic entropies are not sufficient. *J Math Psychol*, in press.
- Sherry CJ, Klemm WR (1982a) Failure of large doses of ethanol to eliminate Markovian serial dependencies in neural spike train intervals. *Int J Neurosci* 17:109–117.
- Sherry CJ, Klemm WR (1982b) Divergence from statistical independence in specified clusters of adjacent neuronal spike train intervals before and after ethanol state changes. *Int J Neurosci* 17:119–128.
- Sherry CJ, Klemm WR (1984) What is the meaningful measure of neuronal spike train activity? *J Neurosci Methods* 10:205–213.
- Stratford KJ, Mason A, Larkman A, Major G, Jack J (1989) The modeling of pyramidal neurones in the visual cortex. In: *The computing neuron* (Durbin R, Miall C, Mitchison G, eds), pp 296–321. Reading, MA: Addison Wesley.
- Strehler BL, Lestienne R (1986) Evidence on precise time-coded symbols and memory of patterns in monkey cortical neuronal spike trains. *Proc Natl Acad Sci USA* 83:9812–9816.
- Tepper JZ, Mandell AJ (1987) Time and frequency domain properties of a cortical functional unit: the discrete penicillin focus in the rat. *Proc NY Acad Sci* 504:301–304.
- Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD (1992) Testing for nonlinearity in time series: the method of surrogate data. *Physica* 58D:77–94.
- Vaadia E, Bergman H, Abeles M (1989) Neuronal activities related to higher brain functions: theoretical and experimental implications. *IEEE Trans Biomed Eng* 36:25–35.
- Villa AEP, Abeles M (1990) Evidence for spatiotemporal firing patterns within the auditory thalamus of the cat. *Brain Res* 509:325–327.
- Wolff JG (1975) Algorithm for the segmentation of an artificial language analog. *Br J Psychol* 66:79–90.
- Wolff JG (1976) Frequency, conceptual structure and pattern recognition. *Br J Psychol* 67:377–390.
- Wolff JG (1977) Discovery of segments in natural language. *Br J Psychol* 68:97–106.
- Wolff JG (1987) Cognitive development as optimisation. In: *Cognitive development as optimisation in computational models of learning* (Bolc L, ed), pp 161–205. New York: Springer.
- Wu X-B, Xu J (1991) Complexity and brain function. In: *Retrospect and prospect of protein research* (Li ZP, ed), pp 247–253. Singapore: World Scientific.
- Wyler AR, Fetz EE, Ward AA (1973) Injury induced long-first-interval bursts in monkey neocortex. *Exp Neurol* 41:773–776.
- Wyler AR, Fetz EE, Ward AA (1975) Firing patterns of epileptic and normal neurons in chronic alumina focus in undrugged monkeys during different behavioral states. *Brain Res* 98:1–20.
- Wyler AR, Burchiel KT, Ward AA (1978) Chronic epileptic foci in monkeys: correlation between seizure frequency and proportion of pacemaker epileptic neurons. *Epilepsia* 19:475–483.
- Wyler AR, Ward AA (1980) Epileptic neurons. In: *Epilepsy: a window to brain mechanisms* (Lockard JS, Ward AA, eds), pp 51–68. New York: Raven.
- Wyler AR, Ojemann GA, Ward AA (1982) Neurons in human epileptic cortex: correlation between unit and EEG activity. *Ann Neurol* 1:301–308.
- Xu J-H, Wu X-B (1992) The information transmission on cerebral cortex based on the complexity measures of EEG signals. In: *Bio-medical modeling and simulation* (Eisenfeld J, Levine DS, Witten M, eds). Amsterdam: Elsevier.
- Zeman W, Innes JRM (1963) *Cragies neuroanatomy of the rat*. New York: Raven.
- Zimmerman ID, Rapp PE, Mees AI (1991) The geometrical characterization of neural activity displays a sensitivity to convulsants. *Int J Bifurcation Chaos* 1:253–259.
- Zvonkin AK, Levin LA (1970) The complexity of finite-objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russ Math Survey* 25:83–124.