Transactions of the American Philological Association 123 (1993) 379-403

Evidence from Experimental Psychology for the Rhythm and Metre of Greek Verse

A. M. Devine
Stanford University

Laurence D. Stephens
University of North Carolina, Chapel Hill

In the traditional philological paradigm of Classical scholarship, questions of theory were largely taken for granted: the theoretical framework of scholarly investigation was thought to be implicit in the textual evidence and so obvious to everyone that it did not merit overt analysis. A highly theoretical topic like the nature of rhythm would have seemed esoteric: certainly it is not a question that the standard handbooks of Greek metre, particularly those with positivist underpinnings, were apt to confront. In recent years, Classics has begun to shed its aversion to theoretical discussion and at the same time has sought a greater degree of integration with a range of related disciplines. So the time has perhaps now come for Classical metrics to access the results of a century of psychological investigation of rhythm.[1] In fact, the a priori exclusion of this information is rather paradoxical, since the Greeks were themselves pioneers in the theoretical study of rhythm and explicitly confronted many of the fundamental problems of the field, as will be illustrated en passant at appropriate points in the ensuing discussion.

1 The internal clock

Greek metres are quintessentially temporal patterns: they develop through time and their constituents are temporally defined; that is, they are patterns *of* time *in* time (Povel 1984).[2] Temporal patterns were one of the early concerns of experimental psychology in the nineteenth century, and although they have

---

[1]We are grateful to *TAPA*'s reviewers, who included an experimental psychologist, for suggestions and revisions; we are responsible for remaining errors.

[2]ῥυθμὸς τοίνυν ἐστὶ σύστημα ἐκ χρόνων κατά τινα τάξιν συγκειμένων. (Aristides Quintilianus 1.13 W[innington-]I[ngram]: "Rhythm is a system of times put together according to some order.")

been intermittently neglected in more recent times, an important body of knowledge has accumulated. Early work was partly motivated by an interest in verse rhythm—for instance, one section of a 1905 article by Stetson is entitled "The rhythm of verse and of prose"—while more recent studies are often conducted from the point of view of the organization of memory or the perception and production of music. Temporal patterns are a subset of serial patterns, having their own specific properties. In temporal patterns, the intervals between the stimuli, and often the stimuli themselves, are temporally defined. Not all temporal sequences of stimuli are patterns; they can be random unpatterned strings. Not all temporal patterns are rhythmic.[3] Rhythm implies some regularly occurring event, although accelerating and decelerating patterns can be learned and reproduced by synchronized tapping with some degree of accuracy (Ehrlich 1958). Although the term rhythm is often used loosely to refer to any regularly occurring event like the succession of day to night or the "rhythm of the seasons," in its technical sense rhythm is often taken to imply a patterned temporal sequence in which the stimuli occur with a frequency within the range of about 8 to 0.5 events per second, or with durations in the range of 120 to 1800 milliseconds (henceforward msec). Slower stimuli tend to be perceived as discrete events not joined to each other in a rhythmic pattern. Faster recurrence leads to various other perceptions such as vibration or tone. In principle, rhythm is not restricted to any one modality.[4] Rhythm is particularly associated with the auditory modality. Coding and reproduction are more accurate for (nonlinguistic) temporal patterns presented auditorily than for similar patterns presented visually (Glenberg et al. 1989), and, more generally, memory for both content and order was better when a list of items was presented auditorily than when it was presented visually (Drewnowski & Murdock 1980). Arrhythmia can be supramodal, discrimination and reproduction being impaired when rhythm is presented via hearing, sight, or touch (Mavlov 1980).

In any motor task—be it speaking, reciting verse, scratching one's head, or running the quarter mile—we need to produce movements whose order is sequenced, whose duration is timed, and whose force is regulated. If someone

---

[3]πιθανὸν μὲν οὖν καὶ χωρὶς λόγου, τὸ μὴ πᾶσαν χρόνων τάξιν εὔρυθμον εἶναι. (Aristoxenus, *E[lementa] R[hythmica]* 8 P[earson]: "It is clear without argument that not every arrangement of times is rhythmical.") οὐ γὰρ πᾶσα χρόνων σύνθεσις εὔρυθμος. (Psellus, *Prolambanomena* 3 P[earson]: "Not every combination of times is rhythmical.")

[4]πᾶς μὲν οὖν ῥυθμὸς τρισὶ τούτοις αἰσθητηρίοις νοεῖται· ὄψει, ὡς ἐν ὀρχήσει· ἀκοῇ, ὡς ἐν μέλει· ἁφῇ, ὡς οἱ τῶν ἀρτηριῶν σφυγμοί· (Aristides Quintilianus 1.13 WI: "All rhythm is perceived by these three sensory modalities: sight, as in dance; hearing, as in song; touch, as the pulse of the arteries.")

is asked to tap his finger at a tempo of his own choice, the tempo chosen will vary greatly from one person to the next, unless they are identical twins, but one and the same person will be surprisingly consistent from one tapping session to another. An interval of about 600 msec between taps is fairly typical. Maximum tapping speed varies according to the part of the body doing the tapping and is in the range 160-205 msec per tap (Keele & Ivry 1987). The phenomenon of spontaneous tapping tempo tells us a number of important things about rhythm. First of all, rhythm is preferred to arrhythmia. In fact, when subjects were asked specifically to tap as irregularly as possible, they found it difficult to do so (Fraisse 1946). Secondly, as just noted, humans are predisposed to perform repetitive motor actions at a specific frequency falling within the range of frequencies perceived as rhythmic. Thirdly, and even more fundamentally, they have the ability to compute that frequency; in other words, spontaneous tapping implies some sort of internal clock.

Not only do people tap with spontaneous rhythm, they can also tap in time to a metronome or march and dance in time to music or work in time to a worksong: in general, this computational ability permits the synchronization of social activity (Shaffer 1982). What is interesting is that when a subject taps to a completely random and unpatterned temporal sequence, the tap follows the signal; just as in most human behavior, the reaction follows the stimulus. But when one taps to a regular series of sounds, the tap is synchronized with the sound (with some degree of variability, which will be discussed below): in fact, the tap occurs about 30 msec prior to the stimulus, suggesting that what is synchronized is the auditory and the tactile perception. Tapping *after* the sound is *more* difficult, particularly when the interval between stimuli is less than 1000 msec (Fraisse 1966; Fraisse & Voillaume 1971). If the experimenter delays one of the stimuli, the tap will occur at its rhythmically predictable point in time. The subject anticipates the stimulus in order to synchronize his tap with the stimulus. He has learned the pattern and is using it to perform the task of synchronization.

There is considerable evidence that humans can efficiently perform only one rhythmic motor task at a time. When subjects pressed one telegraph key with the left hand and another with the right hand in response to two different rhythms, the respective tapping rhythms interfered with one another unless they were harmonically related (Klapp 1979). A similar interference was found when tapping and repeating the syllable *la* to different rhythms (Klapp 1981), or when tapping and reciting a nursery rhyme (Peters 1977). A related difficulty was found with the perceptual monitoring of two temporal sequences (Klapp et al. 1985).

2 The metrical elements

Perhaps the most basic question in Greek metrical theory is how to define the metrical elements. Since Greek metre is quantitative, many theorists have automatically assumed that the longum and breve are temporally (quantitatively) defined, as their names imply. This leads immediately to the question of what specific temporal relations are involved—absolute durations, a temporal ratio, a range of durations or of ratios? Other theorists have taken the position that it is inappropriate to seek either a temporal or a proportional definition for the metrical elements: implicit in their position is a more abstract, nontemporal definition of the elements as strong and weak or plus and minus; this view implies that the metrical elements are encoded and stored in memory without any specifically temporal features, and, a separate point, not in terms of any clearly defined quantitative value or ratio. The question is not whether temporal patterns can be represented at this level of abstraction—they clearly can, since a temporal pattern can routinely be translated into a spatial pattern; rather, the question is whether this level of abstraction is the normal way for temporal patterns to be processed or whether it is only accessed for special purposes. It is important to realize that defining the metrical elements abstractly as strong and weak does not eliminate the question of the nature of the temporal relations of the elements: it just shifts it from that part of the metrical processor which defines the intrinsic features of the elements to that part which specifies the rules for mapping categories of language or speech onto the metrical elements. To put it more concretely, if longum is not temporally defined, then some sort of temporal quantification will be required in the rules specifying what sorts of syllables can implement the longum in acceptable verses: how do the syllables that can implement the longum differ from those that cannot? Just the same possibilities present themselves at this point as those noted above for the other theory: absolute duration, temporal ratio, or range of durations or ratios? So the implication of representing the pattern elements abstractly as strong and weak would be that these temporal pattern elements do not share the durational properties of the linguistic categories that implement them.

In the following exposition, we shall take the most straightforward view, namely, that the metrical elements are temporally defined; the point of the above remarks was to argue that the temporal properties of rhythm are basic to any discussion of quantitative verse irrespective of the position adopted in this particular theoretical dispute and irrespective of the formalism chosen to represent the metrical elements in the medium of typography.

It is fairly obvious that absolute durations are not used to define temporal pattern intervals. If they were, the same pattern would be judged to be two different patterns at two slightly different rates of delivery.[5] It follows that the intervals of temporal patterns are not simply perceived and stored as strings of absolute durations, but are represented in terms of one another, that is, relationally.[6] Pattern elements are represented in our minds in more abstract terms than the physical measurement of their occurrence in our natural environment. Consequently, we need to know what properties characterize and constrain this abstract representation. Are all relationships rhythmically satisfactory? If not, which ones are better than the others? For instance, are ratios like 1:2 (.5) and 1:3 (.33) better than ratios like 2:5 (.4) and 3:4 (.75)? Is there a certain range of ratios that is better than other ranges of ratios? Is there a limit on the number of different ratios that can be efficiently processed as defining different pattern elements? The Greek tradition postulated a basic clock unit that could be used in increments of 100% only;[7] the ratio of 3:2 was admitted for subconstituents of the foot but not as a clock division.[8] It was demonstrated

[5]εἴπερ εἰσιν ἑκάστου τῶν ῥυθμῶν ἀγωγαὶ ἄπειροι, ἄπειροι ἔσονται καὶ οἱ πρῶτοι [scil. χρόνοι]. (Aristoxenus, *Protos Chronos* P: "If the tempi of each rhythm are infinite, then the primary time units will also be infinite.")

[6]οὔτε γὰρ πόδας συντίθεμεν ἐκ χρόνων ἀπείρων, ἀλλ' ἐξ ὡρισμένων καὶ πεπερασμένων μεγέθει τε καὶ ἀριθμῷ καὶ τῇ πρὸς ἀλλήλους ξυμμετρίᾳ τε καὶ τάξει, οὔτε ῥυθμὸν οὐδένα τοιοῦτον ὁρῶμεν· δῆλον δέ, εἴπερ μηδὲ πόδα, οὐδε ῥυθμόν, ἐπειδὴ πάντες οἱ ῥυθμοὶ ἐκ ποδῶν τινων σύγκεινται. καθόλου δὴ νοητέον, ὃς ἂν ληφθῇ τῶν ῥυθμῶν, ὅμοιον εἰπεῖν ὁ τροχαῖος, ἐπὶ τῇσδέ τινος ἀγωγῆς τεθεὶς ἀπείρων ἐκείνων πρώτων ἕνα τινὰ λήψεται εἰς αὐτόν. (Ibid.: "We do not compose feet from infinite (indefinite) time units, but rather from time units that are definite and complete in both magnitude and number and in their commensurability, one to another, and in their order, nor do we see any such (infinite) rhythm. It is clear that if we do not have such a foot, we do not have such a rhythm, since all the rhythms ar composed of certain feet. In general, it must be realized that whatever rhythm be chosen, e.g., the trochaic, when it is set at a particular tempo, it will take for itself a certain one of the infinte number of primary time units.")

[7]καλείσθω δὲ πρῶτος μὲν τῶν χρόνων ὁ ὑπὸ μηδενὸς τῶν ῥυθμιζομένων δυνατὸς ὢν διαιρεθῆναι, καταμετρούμενος, τρίσημος δὲ ὁ τρίς, τετράσημος δὲ ὁ τετράκις. (Aristoxenus, *ER* 10P: "Let the term primary time unit designate the one unit of time lengths that cannot be divided by what is to be rhythmicized; let the term diseme designate the unit that is measure twice as long as this, and the term triseme the unit three times as long.")

[8]τῶν ποδικῶν λόγων εὐφυέστατοί εἰσιν οἱ τρεῖς, ὅ τε τοῦ ἴσου καὶ ὁ τοῦ διπλασίου καὶ ὁ τοῦ ἡμιολίου. (Psellus, *Prolambanomena* 9P: "The most natural of the ratios within feet are the three following: the equal [1:1], the double [2:1], and the hemiolic [3:2].")

almost half a century ago that when subjects are asked to produce patterns of five or six taps having an interval structure of their choice the durations of the intervals used fell into two categories only, one in the range of 200–300 msec and the other in the range of 450–900 msec (Fraisse 1946, 1956). A similar result was obtained when subjects were asked to reproduce patterns having unequal feet, which were described to them on a card in numerical representation (for instance, 322 representing a group of three taps followed by two groups of two taps). Each pattern produced used two intervals, a short interval averaging 320 msec for foot internal taps and a long interval averaging 645 msec for foot final taps; sometimes a dummy tap movement was inserted in the middle of the long interval (Essens & Povel 1985).

Another method used was to get the subject to imitate an auditory pattern by tapping. For instance, in one set of experiments, the stimuli were sequences of 150 msec beeps separated by varying intervals, which the subjects imitated by tapping on a small metal plate which likewise produced a beep (Povel 1981). Two patterns were used, ⏑— and ⏑— —, with the following ratios of intervals between the onsets of adjacent nonidentical tones: .25 (1:4), .33 (1:3), .4 (2:5), .5 (1:2), .6 (3:5), .66 (2:3), .75 (3:4), .8 (4:5). Only patterns in which the intervals stood in the relation 1:2 were correctly imitated. Furthermore, the errors in reproduction of the other ratios were not random: there was a systematic tendency for responses to move closer to a 1:2 ratio. For instance, a stimulus ratio of .25 would be reproduced as .33 and a stimulus ratio of .66 as .55. The subject is evidently aware that the ratio is not .5, but he underestimates or overestimates it in the direction of .5; he strives to represent complex temporal relationships in terms of a simple 1:2 metrical structure. Note that, whereas an interval of 250 msec was not perceived as a subdivision of a 750 msec interval in the sequence 250 750 250 750, it was so perceived (and, consequently, accurately reproduced) in the sequence 250 250 250 750. These results confirm that pattern intervals are not encoded in terms of absolute durations but relationally, and indicate that relations other than 1:2 are not favored.

In another study (Deutsch 1986), subjects were presented with a standard duration demarcated by a pair of blips (50 msec 1 kHz tones) followed after a certain interval by a comparison duration also defined by a pair of blips. The task was to judge whether the comparison duration was equal to, longer than, or shorter than the standard. Additional blips were interpolated in the interstimulus interval and the subjects were instructed to ignore them. However, when the interpolated intervals were somewhat shorter or longer than the standard, their judgment was distorted in the direction of the interpolated durations, but when they were in the range of half the standard duration, their judgment was distorted in the direction of *twice* the interpolated durations. Subjects' judgments were based on a metrical analysis and not on actual durations.

Music uses a much more complex system of durations than verse, but in one study of a variety of piano pieces it was found that in most cases two notes only, standing in a 1:2 ratio, accounted for over 80% of all notes and that among the remaining notes the longer ones tended to be demarcative (Fraisse 1956). When musically trained subjects listened to short musical sequences in which the durations of two notes were systematically varied, they tended to perceive the varying ratios categorically as 1:1 or 2:1 (Clarke 1987). When some of a group of musically trained subjects were asked to play intermixed doublets (two notes per beat) and triplets (three notes per beat) in time to a metronome, they were unable to play the triplets and reorganized them binarily as two sixteenth notes followed by an eighth note (Vos & Handel 1987).

What all this evidence suggests is that longum is discriminated from breve on the basis of a clock defined in terms of breve, and that there is some sort of psychophysical predisposition to a 1:2 ratio which may be expected to apply to quantitative metrical patterns too, since they are a type of temporal pattern. This predisposition was also found to apply to visual patterns when subjects graphically reproduced tachistoscopically presented rows of dots separated by different intervals (Fraisse 1956).

It should also be pointed out that the tendency toward a .5 ratio is not directly related to pure discriminability. There is no reason to assume that long intervals exceeding 200% of the clock interval should be more difficult to discriminate than intervals of 200%. As for intervals below 200%, it is traditionally assumed that intervals of roughly 110% of the standard are just noticeable, and the same holds for continuous stimuli like unseparated tones (Woodrow 1951; Lehiste 1970; Abel 1972); this leaves a range of 90% (110% to 200%) unaccounted for. The just-noticeable delay in a pattern of otherwise isochronous tones was about 6% for intertone intervals of 200 msec and about 12% for intervals of 100 msec (Hirsch et al. 1990). Evidently, quite small differences in duration can be heard, but this ability is not what is used in recognizing temporal pattern intervals as same or different. This latter task seems to be approached in terms of assessing whether the ratio of one interval to the next involves one or two units of an internally computed metrical clock, what are called morae in phonology. When the experimental design forces subjects to compute ratios of duration, the just noticeable differences cited above are not found. Subjects were able to distinguish iambic sequences consisting of noise bursts with the ratios of 1:2 and 2:3 from trochaic sequences with the ratios 3:2 and 2:1, but were not able to distinguish one type of iambic or trochaic sequence from the other. This was true whether the native language of the subjects was English or Estonian, which has ternary duration (Fox et al. 1987).

Although the point will be amplified below, it is worth remarking already at this juncture that both in the simple tone patterns of these psychological experiments and in the far more complex prosody of speech, the definitional 1:2 ratio is subject to modification reflecting the higher level hierarchical structure of the pattern, specifically the grouping of the pattern elements into rising or falling rhythms and the demarcation of higher level structures. These modifications need to be factored out in evaluating the accuracy not only of 1:2 ratios but also off 1:1 ratios in actual production tokens. Similarly, in music performance, the time values of notes are not precisely respected for a variety of reasons, including accent marking, demarcative decelerando, and expressive factors (Gabrielsson 1974; Bengtsson & Gabrielsson 1983; Clarke 1985). It follows that music is abstractly organized in terms of simple integral ratios and concretely implemented by complex nonintegral ratios. Just like quantitative verse.

3 Anceps

So far we have proceeded on the assumption that longum and breve are the only two elements of Greek stichic metre: anceps is interpreted as a position in the metron, namely, the noninternal arsis position, which can be filled by either of the two elements. There has been some controversy about the status of anceps: it is sometimes assumed that anceps is not merely a position but a separate third metrical element with a temporal value intermediate between that of breve and that of longum, thereby imposing a metrical rather than a rhythmical interpretation on the ancient doctrine of *alogia*.[9] We have discussed this question elsewhere (Devine & Stephens 1975).

---

[9] εἰ ληφθείησαν δύο πόδες, ὁ μὲν ἴσον τὸ ἄνω τῷ κάτω ἔχων καὶ δίσημον ἑκάτερον, ὁ δὲ τὸ μὲν κάτω δίσημον, τὸ δὲ ἄνω ἥμισυ, τρίτος δέ τις ληφθείη πούς παρὰ τούτους, τὴν μὲν βάσιν ἴσην αὐτοις ἀμφοτέροις ἔχων, τὴν δὲ ἄρσιν μέσον μέγεθος ἔχουσαν τῶν ἄρσεων. ὁ γὰρ τοιοῦτος πoὺς ἄλογον μὲν ἕξει τὸ ἄνω πρὸς τὸ κάτω· ἔσται δ' ἡ ἀλογία μεταξὺ δύο λόγων γνωρίμων τῇ αἰσθήσει, τοῦ τε ἴσου καὶ τοῦ διπλασίου. καλεῖται δ' οὗτος χορεῖος ἄλογος. (Aristoxenus, *ER* 20P: "If two feet are chosen such that the first has its arsis equal to its thesis and each a diseme in length, and the second has its thesis of diseme

Our concern in the context of this paper is to evaluate the interpretation of anceps as a third metrical element in the light of what is known about temporal patterns in general. Note that in the following discussion reference is always to the abstract durational pattern of the iambic or trochaic metron defined in terms of the putative metrical elements breve, anceps, and longum, and not to individual instantiations of that pattern containing so called long or short anceps, that is, heavy or light syllables in anceps position. Two conditions can be envisaged. In one condition, which is the one envisaged by Aristoxenus in the passage cited in note 9, the temporal value of the longum in the foot containing the anceps is the same as it is in the foot containing the breve, so that the two feet of the metron are anisochronous. In the other condition, the duration of the longum in the foot containing the anceps is compensatorily adjusted so that the two feet of the metron are isochronous (no philological evidence has been adduced that would indicate a preference for shorter heavy syllables in the longum either of feet containing the anceps element or of feet which have heavy syllables in anceps).

There is one recent experiment that seems relevant to the anisochronous condition (Essens et al. 1985). In this experiment, subjects were asked to reproduce patterns of three anisochronous feet made up of tones and intervals, type ∪∪— | ∪— | ∪—, in which the ratio of the breve to the longum was varied through the range 1.5:1, 2:1, 2.5:1, 3:1, 3.5:1, 4:1. The patterns were imitated most accurately when the ratio was 2:1; ratios higher than 2:1 were underestimated in reproduction; the ratio 1.5:1, which is the one that most closely approximates the sort of ratio posited for anceps to breve, was reproduced as almost 2:1; subjects found it difficult to segment patterns with this ratio into feet and in their reproductions they lengthened the longer interval and shortened the shorter one. These results do not encourage the idea that elements standing in a ratio of 1.5:1 are very natural or very easy to discriminate in a temporal pattern of this type. However, the patterns used did not match the metrical distribution of anceps in that only two values were used in the patterns, whereas an iambic or trochaic metron would contain three elements—breve, anceps, and longum—if anceps were a metrical element. This problem does not occur in the experiment relevant to the isochronous condition

length and arsis half of that, and if some third foot is chosen alongside them, such that its thesis is equal to both of the others but its arsis has a magnitude in the middle of the two other arses; then such a foot will have an irrational ratio of arsis to thesis. The irrational will be midway between two ratios perceptible to the senses, i.e., the equal [1:1] and the double [2:1]. This is the irrational choree.")

(Essens 1986 amplifying Povel 1981). Here patterns contained four feet; in one set of patterns, one foot was divided $^2/_3 : {}^1/_3$, while another foot was varyingly divided $^1/_3 : {}^1/_3 : {}^1/_3$ or $^1/_2 : {}^1/_2$. For instance, a structure having trimoraic feet $^1/_3\ {}^1/_3\ {}^1/_3 \mid 1 \mid {}^2/_3\ {}^1/_3 \mid 1$ would be tested against a structure $^1/_2\ {}^1/_2 \mid 1 \mid {}^2/_3\ {}^1/_3 \mid 1$. The $^1/_2$ value in the latter pattern is intermediate between $^1/_3$ and $^2/_3$ and splits the mora, giving a ternary division 1:1.5:2, which is comparable to that posited for breve, anceps, and longum. In the experiment, the patterns with the $^1/_2$ valued intervals were not reproduced accurately either by subjects with extensive musical training or by subjects having no musical training. The pattern was not reanalyzed in terms of the nonoccurring lowest common denominator $^1/_6$.

In musical performance, dotted eighth notes tend to be lengthened and dotted sixteenth notes to be shortened (Gabrielsson 1974); recategorization of dotted notes was particularly associated with faster tempi (Clarke 1985). A fairly rapid twelve-beat sequence divided unevenly (for instance, 22323) is found in African lyrics (Jones 1964), and similar structures occur in Macedonian and Bulgarian folk songs (Singer 1974; Pressing 1983); in such sequences the divisions, as evidenced by hand clapping or the downward segment of dance movements, fall at anisochronous intervals, but there is no unit intermediate between one mora and two such as that posited by the anceps theory.

Despite the popularity enjoyed by the theory of anceps as a metrical element, we find it difficult to believe that large tracts of Greek verse were written in patterns using an element which the performer would have difficulty generating and which the audience would have difficulty perceiving.

4 Feet[10]

4.1 Chunking

One of the best known and most fundamental characteristics of the human mind is its drive to relate and organize the information it is processing. For instance, subjects were presented with a list of sixty items belonging to four semantic

[10] ᾧ δὲ σημαινόμεθα τὸν ῥυθμὸν καὶ γνώριμον ποιοῦμεν τῇ αἰσθήσει, πούς ἐστιν εἷς ἢ πλείους ἑνός. (Aristoxenus, *ER* 16P: "That by which we signal rhythm and render it perceptible to the sense is the foot, either one or more than one.") καὶ ἔστι ῥυθμὸς μὲν ὥσπερ εἴρηται σύστημά τι συγκείμενον ἐκ τῶν ποδικῶν χρόνων ὧν ὁ μὲν ἄρσεως, ὁ δὲ βάσεως, ὁ δὲ ὅλου ποδός. (Psellus, *Prolambanomena* 8P: "Rhythm, as just stated, is a system composed of podic temporal units, of which one is that of the arsis, another that of the thesis, and another that of the whole foot.") πoὺς μὲν οὖν ἐστι μέρος τοῦ παντὸς ῥυθμοῦ δι' οὗ τὸν ὅλον καταλαμβάνομεν. (Aristides Quintilianus 1.14WI: "The foot is a part of every rhythm through which we comprehend the whole.")

categories, e.g., animals, names, professions, and vegetables; the items in the list were in random order; when asked to recall the items, subjects reproduced them not in random order but in categorial clusters (Bousfield 1953). When a list is learned incrementally, with items presented one at a time, recall starts in the order of presentation, but after a few items have been presented, it is restructured into semantically categorial clusters (Mandler & Dean 1969). Categorial organization is not limited to relating items to each other at a single level by classifying them into equipollent categories: categorial structure can also be hierarchical. Recall was found to be better for a list of forty words consisting of four supercategories each containing two categories of five items each than for a list of forty words consisting of four categories of ten items each (Cohen & Bousfield 1956).

In other tasks, particularly those using strings of digits or letters, the subject is required to remember not only the number and content of the items but also their serial order; remembering telephone numbers is a comparable everyday activity. Because the error patterns for order-only recall were different from those for item-only recall, and because the former is independent of the memory load for the latter, it has been assumed that the two types of information are stored separately (Healy 1974, 1982). Nevertheless, changing order information can be just as grave as changing item information (Johnson 1970). The categories used to organize serial information are serial segments, or chunks as they are called. Interest in chunking was stimulated by the observation that, while the capacity of short-term memory was notoriously limited, it could be significantly increased if sequences of items were recoded into informationally richer chunks. Memory capacity then depended more on the number of chunks than on the total number of items in the sequence. For instance, when subjects were trained to recode strings of ones and zeros into digits representing two items each, e.g., 01→1, 10→2, 11→3, recall improved even more (Miller 1956). The tradeoff is that the recoding rules have to be learned and used.

There are various indications that a sequence has been recoded into chunks. After learning lists of twelve nonsense syllables that he organized into six accentual trochees, the subject learned a reordered list with the original feet preserved much more rapidly than one consisting of contiguous syllables across the foot boundary, indicating that the syllables had been coded as nonsense words (Mueller & Schumann 1894). A study of response times in alphabetic retrieval found that subjects chunked the alphabet in the same way as it is phrased in the alphabet songs taught in the nursery schools: they accessed the earlier items of a chunk more rapidly than later items; this differential access suggests that a chunk exists and that it is unpacked from memory starting at the beginning (Klahr et al. 1983). When asked if a pair of letters was part of a previously memorized letter sequence, subjects answered more rapidly if both letters belonged to the same chunk than if they straddled a chunk boundary (Johnson 1978). Error patterns in recall are quite different for a chunked sequence and an unstructured serially associated sequence. In the latter, errors tend to cluster in the middle of a sequence; in the former they cluster at the beginning of chunks. Recall of the first item of a chunk implies recall of the following item in the same chunk more than recall of the last item in a chunk implies recall of the first item of the following chunk (Johnson 1970). Error probabilities at chunk seams are termed transition(al) error probabilities. Pause patterns in spoken recall are another indication of chunking: when subjects read and memorized groups of letters on file cards, in their spoken recall pausing between groups was longer than pausing within groups both in forward and in backward recitation of the list (McLean & Gregg 1967). Conversely, subjects report perceiving pauses at the end of accentually dactylic feet even though the stimuli were presented at a constant rate (Jones 1978). Pause in auditory presentation and space in visual presentation can be used by the experimenter to induce subjects to chunk a sequence in some particular way. The resulting chunk organization is an intrinsic part of the learned sequence: learning the letter sequence SBJ FQLZ was no faster for subjects who had previously learned SB JFQ LZ than it was for subjects who had previously learned an entirely different sequence (Johnson & Migdoll 1971). Similarly, a second presentation of a string of digits with different grouping was not recognized as a repetition of its earlier occurrence (Bower & Winzenz 1969); and the same effect was found with nonverbal material, namely, minimelodies of five tones each (not in key) (Dowling 1973).

The optimum length for a chunk has also been studied.[11] People find it difficult to segment sequential information into chunks larger than three or four items (Estes 1972); pausing patterns in free recall experiments also point to chunks of three or four items (Broadbent 1975). When subjects were presented with a list of digits and instructed to rehearse them in groups of varying sizes, recall improved as group size grew from one to two to three and deteriorated as it further increased to four and five; deterioration for groups larger

[11] οὐ καθ' αὑτὸν ὁ ποὺς εἰς τὸ πλέον τοῦ εἰρημένου πλήθους [sc. τῶν τεττάρων ἀριθμῶν] μερίζεται. (Aristoxenus *ER* 19P: "A foot is not, *eo ipso*, divided into more than the number of parts stated [sc. four].")

than three affected serial order rather than item recall, suggesting that constraints on chunk size may be related to memory for order more than for item (Wickelgren 1964, 1967). Another study found, on the basis of transition error analysis, that four-item sequences were treated as single chunks by half the subjects and as two chunks by the other half, and that five-item sequences were almost invariably analyzed as either 2+3 or 3+2 (Johnson 1970).

4.2 Subjective rhythm

The letter and digit sequences that are used in many chunk studies are revealing because there are few potentially confounding objective properties of the stimulus to disturb the subject's choice of chunk size or reaction to grouping imposed by the experimenter. When the stimulus sequence is patterned, as it is in verse, the properties of the pattern are a primary factor conditioning grouping. In fact, pattern induction is such a basic cognitive activity that we have a propensity to look for and "find" patterns even when they are objectively not there. When presented with randomly sequenced binary events, subjects behave as though the stimuli were patterned and predict grouped subsequences of events (Simon & Sumner 1968). Our perception of isochronous auditory events having a frequency within the range of rhythmical perception tends not to be a string of equipollent elements; rather, we have the impression that the sequence is grouped into subsequences of two or three. The most familiar instance is the tick-tock (not, significantly, tick-tick) of a clock. The tick tends to give the impression of having greater intensity and less duration ("i" is often sound-symbolic of the diminutive) and the tock of having greater duration or of being followed by a longer silent interval, and the tick and the tock together give the impression of forming a group that is separate from preceding and following events. This perceptual characteristic may explain why alternating rhythm has been found when subjects tap in time with an isochronous tone and continue this synchronized tapping after the tone ceases. Finger taps produced under these conditions are not strictly isochronous; rather, there is variation such that longer taps tend to be either preceded or followed by shorter taps. This distribution was explained as arising mainly from unintentional variability in the motor execution of programmed isochrony (Wing & Kristofferson 1973a, 1973b; Wing 1980). Even though the variations in the motor execution of the taps are independent of each other, they will produce a nonrandom pairing of long-short or short-long taps in the following way. Two adjacent intertap intervals are affected by three motor delays. The first at the beginning of the first interval and the third at the end of the second interval are uncorrelated in their respective effects, but the second motor delay has opposite but exactly equal effects on the two intervals: if it postpones the end of the first interval longer than the average, it necessarily postpones the beginning of the second longer than the average, and vice versa, so that longer intervals will be followed by shorter ones and vice versa; i.e., adjacent intertap intervals will be negatively correlated. Nonadjacent intertap intervals, however, not being affected by any one motor delay in common, would be uncorrelated. However, there is also an alternation in the force of the taps, and it has recently been suggested that the isochronous tone stimuli are preceived as temporally and accentually differentiated and, consequently, replicated with some degree of alternating rhythm in production; alternating rhythm implies foot structure (Nagasaki 1987a, 1987b, 1990). In experiments performed at the turn of the century, subjects were presented with sequences of twelve nonsense syllables and reproduced the sequence with stress on alternate syllables, grouping the resulting feet into hemistichs. When they suppressed their tendency to impose alternating rhythm on the sequence, the sequence had to be repeated twice as many times in order to be learned accurately (Mueller & Schumann 1894). A recent study of order permutation of sequences of visually presented nonsense syllables also revealed foot structure (Gordon & Meyer 1987).

4.3 Iambic and trochaic foot structure

Up to this point, we have considered the grouping of random sequences like letter sequences and the perceptual grouping of physically isochronous sequences of identical stimuli. However, many temporal patterns are made up of objectively differentiated repeating events. In an auditory pattern, the most obvious differentiating properties are intensity, duration, and frequency (how loud the sound is, how long it lasts, and what its pitch is). The relative contribution of these three parameters to grouping has been studied experimentally since the turn of the century. In one famous early experiment (Woodrow 1909), when subjects were presented with a regular series of sounds lasting 135 msec followed by a silent interval of 615 msec of which every alternate sound had greater intensity, they uniformly perceived trochaic rhythm; that is, the less intense sound was grouped with the more intense sound in such a way that the less intense sound ended the foot. The next step in the experiment was to increase gradually the duration of the silent interval following the more intense sound and correspondingly decrease the duration of the silent interval preceding the intense sound, thereby maintaining a constant clock measure for the foot. For instance, instead of having 615 msec following each sound, you would have 603 msec preceding and 627 msec following, or 547 msec preceding and 683 msec following. As this was done, the perception of trochaic rhythm became progressively weaker until, passing through a neutral stage of ambivalent grouping, it began to change into iambic rhythm. The intermediate neutral stage was particularly interesting, because it could be used to measure the degree of change in duration that was required to counterbalance the contribution of any particular degree of intensity to the grouping of the stimuli. The effect of varying the durations of the sounds and their following silent intervals in the absence of any difference in intensity was also studied both in this early experiment and in more recent ones. In general, the longer the relative duration of the period from the onset of one tone to the onset of the next, the more likely that longer element is to sound accented and to end the group, that is, the greater the likelihood of iambic grouping. Starting with an isochronous sequence generated from 50 msec tones followed by 50 msec intervals, every second interval was gradually increased: subjects began to perceive groups of tones ending in the longer interval but with a comparatively weak accent on the first tone as soon as the longer interval exceeded the shorter interval by 5-10%. When the longer interval was further increased, the perceived grouping remained the same, but a strong accent was now reported on the second tone of the group; it required a fairly robust increase in intensity (4 dB) on the first tone to counterbalance this accent on the second tone and produce a perception of equal accents (Povel & Okkerman 1981).

Another interesting study (Vos 1977) varied both tone duration and silent interval. Half of the subjects were professional musicians and half philologists from the Department of Classics at the University of Nijmegen. The stimuli were strings of thirty pure tones all having the same frequency and the same intensity, except that the intensity was gradually increased from 0 dB to 40 dB over the first two seconds to minimize the orientation effect (see below). The first two tones and the first two following silent intervals were assigned durations of either 80 msec (short) or 320 msec (long), and the resulting pattern was repeated throughout the string. Subjects made a forced choice judgment as to whether the string was trochaic or iambic. When one of the tones was long and the other short, the long tone was judged prominent. When one tone was long and the other short and their intervals were identical (320_{80} 80_{80}), the string was judged iambic, i.e., the long tone was judged to end the foot. But when the long tone was followed by the short interval and the short tone by the long interval (320_{80} 80_{320}), the string was judged trochaic: the longer tone was prominent, but the longer interval demarcated the feet. When both tones were short but one interval was long and the other short (80_{320} 80_{80}), the tone followed by the longer interval was perceived as foot final and so the string was judged iambic. The staccato nature of many of these experimental stimuli (as compared to the relatively legato prosody of speech, in which nondemarcative periods of silence are associated only with the closure portions of stop consonants) is very useful, since by uncoupling tone duration from silent interval duration, it tends to confirm the idea that whereas the strength of the signal cues the accent, the duration of the silent interval cues the end of the group. Consequently, the grouping of temporal patterns into feet is based on the gestalt principle of proximity (Koffka 1962) according to which closely spaced elements are more likely to be grouped together than distantly spaced elements as in the following visual pattern:

❒ ❒   ❒ ❒   ❒ ❒   ❒ ❒

When subjects were asked to tap groups of three or four taps, they spontaneously separated the groups by pauses (Fraisse 1982). In another experiment already described above, when subjects were instructed to produce patterns described to them in terms of the number of taps per foot, they spontaneously used short intervals for foot internal taps and long intervals for foot final taps (Essens & Povel 1985). In the synchronized tapping to an isochronous tone described above, when the interval between tone onsets was in the range 250-375 msec, the grouping was accentually alternating with the accented element having longer duration; when the interval was in the range 200-250 msec, the grouping was accentually alternating with the unaccented element having longer duration (Nagasaki 1987a). When subjects were asked to synchronize taps with an isochronous series of sounds in which every other sound had greater intensity, the intervals between the taps varied due to two factors. The more intense sound induced a slightly longer interval, and this effect was considerably more robust when the series was perceived as having iambic grouping: trochees were tapped with 484 msec for the more intense sound and 452 msec for the less intense foot final sound, whereas iambs were tapped with 432 msec for the less intense sound and 520 msec for the more intense foot final sound (Fraisse 1956: 95).

It has also been suggested that foot demarcation in alternating rhythm is reflected in differential electrical activity (evoked potentials) elicited in the brain which can be recorded via electrodes placed on the scalp (care is needed in the interpretation of these results). It was found that the amplitude of event related potentials tended to decrease foot internally but to increase after a foot boundary pause, possibly reflecting a recovery of sensitivity during the pause period (Fraisse & Lavit 1986).

4.4 Orientation effect; run unity

In Greek, the difference between iambic and trochaic in the surface ratios of breve to longum apparently does not reflect a difference in the definition of either the metrical elements or the categories of linguistic syllable that can implement them: longum is implemented by heavy syllables and breve by light syllables in both iambic and trochaic verse, with the possible exception of the occasional responsion of a first paeon sequence with a trochaic metron in comedy, e.g., *Wasps* 1062–1093. Rather, the difference reflects properties of demarcation that are implemented as the elements are processed for production. It follows that, disregarding the effect of differences in the coincidence of word boundary with foot boundary, the crucial distinction between iambic and trochaic at the abstract pattern level must lie in how the pattern starts. It is pointed out in a number of experimental studies (Woodrow 1909; Fraisse & Oléron 1954; Handel & Yoder 1975; Povel & Okkerman 1981) that the pattern structure initially presented to subjects can determine the outcome of the analysis, and various strategies have been adopted to circumvent this orientation effect, such as instructing subjects to try both rising and falling rhythm or prefixing an extrametrical tone to the pattern. The orientation effect was studied with patterns of nine High and Low tones, type HLHHHLLLL, presented in continually repeating sequence with an interonset interval of 333 msec; subjects were more likely to use the starting pattern as their perceived pattern organization than any other analysis (Preusser et al. 1970). This orientation effect was also critical when subjects were asked to classify Dutch recited verses as iambic or trochaic (Loots 1980). One factor that can counteract the orientation principle is the tendency for sequences of like events to be grouped together: this is termed "run unity." In patterns similar to those just described, subjects rarely if ever analyzed the pattern as beginning in the middle of a string of identical sounds (Royer & Garner 1970). In comparatively unpatterned strings of binary events like 110000110, errors clustered at the ends of runs of like events, and recall depended on the number of runs in the sequence rather than on the number of events, indicating that runs were coded as chunks (Keller 1963). Run unity reflects a general gestalt principle that also applies to visual patterns, as in the following:

■ ■ □ □ ■ ■ □ □ □ □ ■ ■ □ □ ■ ■	versus	■ □ ■ □ ■ □ ■ □ ■ □ ■ □ ■ □ ■ □

When taken together, run unity and the orientation effect suggest that pattern perception is a hypothesis testing activity that starts at the beginning of the stimulus and privileges hypotheses that do not require the pattern to start in the middle of a foot. Once a hypothesis is chosen, it is coded into memory: consequently, pattern violations are less serious at the beginning of a pattern than internally. Disruptions to an isochronous tone sequence were more easily detected when the disrupted sequence followed the regular sequence than when it preceded (Bharucha & Pryor 1986). Line initial freedom is a well-known property of verse. The initially chosen hypothesis can be discarded if later information indicates it was inadequate.

In metrical patterns, run unity implies that a series of amphibrachs will tend to be analyzed as dactyls or anapaests even if it begins with an amphibrach, provided the amphibrachs are presented continuously, that is, provided the amphibrach structure is not reinforced by demarcative pauses; the initial amphibrach is interpreted as anacrustic. This ties in with what we know about language and about Greek metre. Languages rarely if ever have amphibrach stress feet (Haraguchi 1991), and Greek does not have amphibrach stichoi. Conversely, when the reiterated demarcative effect of word boundary induces amphibrach grouping in dactylic, the result is a disfavored structure, as in the notorious ἡ δὲ χίμαιραν ἔτικτε πνέουσαν (*Theogony* 322) or the sound symbolic αὖτις ἔπειτα πέδονδε (*Odyssey* 11.598). The amphibrachic patterning of the word boundaries ("pause"-induced grouping) obscures the dactylic patterning of the syllable weights (run induced grouping). One study relevant to this phenomenon took a repeating eight element sequence of alternating tones such as HH LL HL HL; when pauses were inserted after even numbered tones, i.e. at the run induced foot boundaries, subjects almost always recognized the tonal pattern, but when pauses were inserted after every third element, thereby splitting the feet, the tonal pattern was only correctly recognized about half the time (Handel 1973).

4.5 Tempo

The grouping of tones into feet has been observed to depend on tempo in a number of different studies (Bolton 1894; Fraisse 1956; Handel & Oshinsky 1981; Povel 1984); the faster the tempo, the more tones are grouped into a foot.[12] Tempo likewise affects the identification of the beat in music perception (Madsen et al. 1986). When asked to tap to the perceived beat, musically

---

[12] διαφέρουσι δὲ οἱ μείζονες πόδες τῶν ἐλαττόνων ἐν τῷ αὐτῷ γένει ἀγωγῇ. ἔστι δὲ ἀγωγὴ ῥυθμοῦ τῶν αὐτῷ λόγῳ ποδῶν κατὰ μέγεθος διαφορά (*Fragmenta Neapolitana* 15P: "In the same genus, the greater feet differ from the lesser in tempo. The tempo of a rhythm is a difference in magnitude of feet which have the same ratio [sc. of arsis to thesis].")

trained subjects reacted to faster stimuli by tapping at half the presented rate and to slower stimuli by tapping at double the presented rate (Duke 1987 reported in Radocy & Boyle 1988). The processing of relatively rapid tempi having intertap intervals of about 250 msec may involve different mechanisms from the processing of slower tempi having intertap intervals of 500 or 1000 msec, since children younger than four years and a brain damaged subject synchronized well with the former but not with the latter (Kohno et al. 1991); speech syllables at ordinary speech rates fall into the former category.

5 Metron

There was already some indication in the letter chunking studies that subjects were not confining their organization to a single level of coding. They not only coded letter sequences into chunks, but could also recode chunks into superchunks or subchunks. Transition error analysis for sequences like NGVHSB showed that some subjects chunked this NG VHSB, while others split the longer sequence into two chunks giving NG VH-SB (where the hyphen marks the binary division of a chunk into subchunks) rather than into the ternary structure NG VH SB with three equipollent chunks (Johnson 1970). In another study, subjects induced structures like J XC-PM by learning a set of rules for changing the order of structural units of the sequence; they were then able to apply these rules effortlessly to a different sequence of letters, indicating that they had correctly abstracted the hierarchical structure with its metron organization (Keeney 1969). Subjects presented with continuous sequences of evenly spaced identical sounds grouped them not only into feet but also into metra, and judged not only one sound in the foot to be stronger than the other but also one foot in the metron to be stronger than the other (Woodrow 1951). In rapid synchronized tapping at intertap durations below 200 msec, the grouping was into an accentually alternating metron with accented elements accorded longer duration, suggesting that at faster tempi feet encompass more taps and are then binarily subdivided (Nagasaki 1987a). When subjects tapped in time to a sequence of 6 isochronous tones in which the first tone was either raised or lowered in pitch, the fifth and either the first or the second intertap intervals were prolonged (Franěk et al. 1991). The evidence cited in this section is sufficient to establish the metron as a structure available for the processing of serial sequences in general, even if it is not as exhaustively corroborated as the foot.

6 Hemistich and stichos

The great potential power of multilevel hierarchical chunking was pointed out in early theoretical analyses (Mandler 1967). A comparatively long and prima facie complicated sequence can be generated by learning a few rules for unpacking progressively smaller superchunks. These rules can be considered as more flexible, abstract, and "intelligent" than the simple and naive strategy of associating a preceding item with a following one (Restle & Brown 1970). In one experiment, a six–item alphabet consisting of a row of six lights was used to present a 32–event sequence to subjects at a slow rate of about 4 seconds per event:

12122323121223236565545465655454

This sequence is much less intimidating once one perceives its hierarchical structure, which is that of a regular binarily branching tree, just like a tetrameter except that there are six elements rather than two,

12-12 23-23 | 12-12 23-23 || 65-65 54-54 | 65-65 54-54 ||,

and, in fact, subjects were able to anticipate the pattern quite successfully (Restle 1970). Such sequences are learned more easily and with fewer errors than sequences that cannot be generated from a tree structure; and if such a sequence contains at some point an individual item that cannot be generated from a tree structure, that particular item will be difficult to learn and responses will often smooth out the irregularity, particularly toward the end of a subunit. Evidently, instead of storing each event sequentially in memory and recalling an event on the basis of its association with the preceding event, subjects encoded the sequence in terms of a few rules for iterating subpatterns in a hierarchical structure, or, to use metrical terminology, for generating metra out of feet, hemistichs out of metra, and stichoi out of hemistichs. Furthermore, the error rate at the metron boundary was less than at the hemistich boundary, and the error rate foot internally was the lowest of all: this distribution of error rates (lower within any domain than across the boundary between two such domains) indicates that subjects were using a hierarchical analysis and not sequential association or random chunking which would not have given rise to such a structured transition error distribution. One other result of processing the information in this way is that one would not have a tendency to produce sequences of thirty-one events or thirty-four events, as might be the case if the sequence were generated by counting to thirty-two each time or by chunking the sequence in a nonhierarchical sequential fashion. So the hierarchical structure of metrical patterns is also an inbuilt safeguard against errors like nine-foot tetrameters, notwithstanding the odd inadequate experimental subject or inscriptional poetaster.

The same tetrameter type structure was investigated with stimuli speeded up so that each light came on for 300 msec followed by an interval of 300 msec in the allegro presentation and 800 msec in the adagio presentation (Restle 1972). These intervals were then varied in such a way as to emphasize or to counteract the hierarchical tree structure. Pattern induction was improved by the former and hindered by the latter condition. Pauses improved pattern induction by demarcating the constituent structure and not merely by providing more time for processing, since rate of presentation was not critical. Hierarchical interval differentiation was a little better than no differentiation at all and much better than counter-hierarchical interval differentiation.

In finger tapping an eight-element sequence using the index and middle fingers with alternating hands, subjects organized the sequence as a binarily branching tree, since intertap intervals were greater between metra than between feet and greater between feet than foot internally (Rosenbaum et al. 1983); similar results were obtained with other tree structures and other finger combinations (Collard & Povel 1982).

Concluding Remarks

What emerges quite clearly from the above review is that the structure of verse patterns is not unique to verse at all, but reflects very general properities of the psychological processing of patterns and particularly of rhythmical patterns. This leaves unanswered a further question, namely: how does verse become patterned? A simple preliminary hypothesis would be that the poet takes some unpatterned raw material—language—and arranges it into patterns according to basic general principles of pattern structure. Parts of the discussion above have been cast in these terms in order to simplify the exposition. So we cannot end without pointing out that such a hypothesis is deficient because it implicitly assumes that language is unpatterned and only becomes patterned in verse. In fact, language and speech themselves involve highly patterned prosodic structures. The more we find out about the prosody of Greek speech, the clearer it is becoming that the constituents of verse structure—syllables, feet, metre, hemistichs, stichoi—are simply more highly constrained, regularized or prototypical instances of prosodic constituents and domains that pre-exist in the Greek language. Verse entails not the arrangement of language into non-linguistic patterns but the constrained choice of the most regular instances of prosodic patterns and structures that occur systematically and naturally in language. So in relating the information reviewed in this paper to verse structure, we need to bear in mind that verse is not the creation of patterns out of language but a regularization of the patterns in language.

Works Cited

Abel, S. 1972. "Duration Discrimination of Noise and Tone Bursts." *Journal of the Acoustical Society of America* 52: 1219.

Bengtsson, I., and A. Gabrielsson. 1983. "Analysis and Synthesis of Musical Rhythm." Studies of Musical Performance, ed. J. Sundberg (Stockholm): 27-60. Royal Swedish Academy of Music Publications 39.

Bharucha, J. J., and J. H. Pryor. 1986. "Disrupting the Isochrony of Underlying Rhythm." *Perception and Psychophysics* 40: 137-141.

Bolton, T. 1894. "Rhythm." *American Journal of Psychology* 6: 145-238.

Bousfield, W. A. 1953. "The Occurrence of Clustering in the Recall of Randomly Arranged Associates." *Journal of General Psychology* 49: 229-240.

Bower, G. H., and D. Winzenz. 1969. "Group Structure, Coding, and Memory for Digit Series." *Journal of Experimental Psychology* Monograph 80.2.2.

Broadbent, D. E. 1975. "The Magic Number Seven after Fifteen Years." *Studies in Long-Term Memory*, ed. A. Kennedy and A. Wilkes (New York) 3-18.

Clarke, E. F. 1985. "Structure and Expression in Rhythmic Performance." *Musical Structure and Cognition*, ed. P. Howell, I. Cross, and R. West (London) 209-236.

Clarke, E. F. 1987. "Categorical Rhythm Perception: An Ecological Perspective." *Action and Perception in Rhythm and Music*, ed. A. Gabrielsson (Stockholm) 19-33. Royal Swedish Academy of Music Publications 55.

Cohen, B. H., and W. A. Bousfield. 1956. "The Effects of a Dual-Level Stimulus-Word List on the Occurrence of Clustering in Recall." *Journal of General Psychology* 55: 51-58.

Collard, R., and D.-J. Povel. 1982. "Theory of Serial Pattern Production: Tree Traversals." *Psychological Review* 89:693-707.

Deutsch, D. 1986. "Recognition of Durations Embedded in Temporal Patterns." *Perception and Psychophysics* 39: 179-186.

Devine, A. M., and L. D. Stephens. 1975. "Anceps." *Greek, Roman, and Byzantine Studies* 16: 197-215.

Dowling, W. J. 1973. "Rhythmic Groups and Subjective Chunks in Memory for Melodies." *Perception and Psychophysics* 14: 37-40.

Drewnowski, A., and B. B. Murdock Jr. 1980. "The Role of Auditory Features in Memory Span for Words." *Journal of Experimental Psychology: Human Learning and Memory* 6: 319-332.

Duke, R. A. 1987. "Musicians' Perception of Beat in Monotonic Stimuli." Reported in R. E. Radocy and J. D. Boyle, *Psychological Foundations of Musical Behavior*, 2nd ed. (Springfield).

Ehrlich, S. 1958. "Le mécanisme de la synchronisation sensorimotrice." *L'Année Psychologique* 58: 7-23.

Essens, P. 1986. "Hierarchical Organization of Temporal Patterns." *Perception and Psychophysics* 40: 69-73.

Essens, P. J., and D.-J. Povel. 1985. "Metrical and Nonmetrical Representations of Temporal Patterns." *Perception and Psychophysics* 37: 1-7.

Estes, W. K. 1972. "An Associative Basis for Coding and Organization in Memory." *Coding Processes in Human Memory*, ed. A. W. Melton and E. Martin (Washington, D.C.) 161-190.

Fox, R. A., and I. Lehiste. 1987. "Discrimination of Duration Ratios by Native English and Estonian Listeners." *Journal of Phonetics* 15: 349-363.

Fraisse, P. 1946-1947. "Mouvements rythmiques et arythmiques." *L'Année Psychologique* 47-48: 11-21.

Fraisse, P. 1956. *Les structures rhythmiques: Étude psychologistique*. (Louvain).

Fraisse, P. 1966. "L'anticipation de stimulus rythmiques." *L'Année Psychologique* 66: 15-36.

Fraisse, P. 1982. "Rhythm and Tempo." *The Psychology of Music*, ed. D. Deutsch (New York) 149-180.

Fraisse, P., and A. Lavit. 1986. "Auditory Evoked Potentials and the Perception of Rhythms." *International Journal of Psychophysiology* 4: 85-89.

Fraisse, P., and G. Oléron. 1954. "La Structuration intensive des rythmes." *L'Année Psychologique* 54: 35-52.

Fraisse, P., and C. Voillaume. 1971. "Les repères du sujet dans la synchronisation et dans la pseudo-synchronisation." *L'Année Psychologique* 71: 359-369.

Franěk, M., J. Mates, T. Radil, K. Beck, and E. Pöppel. 1991. "Sensorimotor Synchronization: Motor Responses to Regular Auditory Patterns." *Perception and Psychophysics* 49: 509-516.

Gabrielsson, A. 1974. "Performance of Rhythm Patterns." *Scandinavian Journal of Psychology* 15: 63-72.

Glenberg, A. M., S. Mann, L. Altman, T. Forman, and S. Procise. 1989. "Modality Effects in the Coding and Reproduction of Rhythms." *Memory and Cognition* 17: 373-383.

Gordon, P. C., and D. E. Meyer. 1987. "Control of Serial Order in Rapidly Spoken Syllable Sequences." *Journal of Memory and Language* 26: 300-321.

Handel, S. 1973. "Temporal Segmentation of Repeating Auditory Patterns." *Journal of Experimental Psychology* 101: 46-54.

Handel, S., and J. S. Oshinsky. 1981. "The Meter of Syncopated Auditory Polyrhythms." *Perception and Psychophysics* 30: 1-9.

Handel, S., and D. Yoder. 1975. "The Effects of Intensity and Interval Rhythm on the Perception of Auditory and Visual Temporal Patterns." *Quarterly Journal of Experimental Psychology* 27: 111-122.

Haraguchi, S. 1991. *A Theory of Stress and Accent* (Dordrecht).

Healy, A. F. 1974. "Separating item from Order Information in Short-Term Memory." *Journal of Verbal Learning and Verbal Behavior* 13: 644-655.

Healy, A. F. 1982. "Short-Term Memory for Order Information." *The Psychology of Learning and Motivation* 16: 191.

Hirsch, I. K., C. B. Monahan, K. W. Grant, and P. G. Singh. 1990. "Studies in Auditory Timing: 1. Simple Patterns." *Perception and Psychophysics* 47: 215-226.

Johnson, N. F. 1970. "The Role of Chunking and Organization in the Process of Recall." *The Psychology of Learning and Motivation* 4: 171-247.

Johnson, N. F. 1978. "The Memorial Structure of Organized Sequences." *Memory and Cognition* 6: 233-239.

Johnson, N. F., and D. M. Migdoll. 1971. "Transfer and Retroaction under Conditions of Changed Organization." *Cognitive Psychology* 2: 229-237.

Jones, A. M. 1964. "African Metrical Lyrics." *African Music* 3.3: 6.

Jones, M. R. 1978. "Auditory Patterns: Studies in the Perception of Structure." *Handbook of Perception* 8 (New York) 255-288.

Keele, S. W., and R. I. Ivry. 1987. "Modular Analysis of Timing in Motor Skill." *The Psychology of Learning and Motivation* 21: 183-228.

Keeney, T. J. 1969. "Permutation Transformations on Phrase Structures in Letter Sequences." *Journal of Experimental Psychology* 82: 28-33.

Keller, L. 1963. "Run Structure and the Learning of Periodic Sequences." Ph.D. Diss. Indiana University, reported in Restle 1967.

Klahr, D., W. G. Chase, and E. A. Lovelace. 1983. "Structure and Process in Alphabetic Retrieval." *Journal of Experimental Psychology: Learning, Memory and Cognition* 9: 462-477.

Klapp, S. T. 1979. "Doing Two Things at Once: The Role of Temporal Compatibility." *Memory and Cognition* 7: 375-381.

Klapp, S. 1981. "Temporal Compatibility in Dual Motor Tasks." *Memory and Cognition* 9: 398-401.

Klapp, S., M. Hill, J. Tyler, Z. Martin, R. Jagacinski, and M. Jones. 1985. "On marching to Two Different Drummers." *Journal of Experimental Psychology: Human Perception and Performance* 11: 814-828.

Koffka, K. 1962. *Principles of Gestalt Psychology* (London).

Kohno, M., A. Kashiwagi, and T. Kashiwagi. 1991. "Two Processing Mechanisms in Rhythm Perception." *Proceedings of the XIIth International Congress of Phonetic Sciences* 2: 86-89. Aix-en-Provence.

Lehiste, I. 1970. *Suprasegmentals* (Cambridge, MA.).

Loots, M. E. 1980. *Metrical Myths* (The Hague).

Madsen, C. K., R. A. Duke, and J. M. Geringer. 1986. "The Effect of Speed Alterations on Tempo Note Selection." *Journal of Research in Music Education* 34: 101-110.

Mandler, G. 1967. "Organization and Memory." *The Psychology of Learning and Motivation* 1: 327-372.

Mandler, G., and P. J. Dean. 1969. "Seriation: Development of Serial Order in Free Recall." *Journal of Experimental Psychology* 81: 207-215.

Mavlov, L. 1980. "Amusia Due to Rhythm in a Musician with Left Hemisphere Damage." *Cortex* 16: 331-338.

McLean, R. S., and L. W. Gregg. 1967. "Effects of Induced Chunking on Temporal Aspects of Serial Recitation." *Journal of Experimental Psychology* 74: 455-459.

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two." *Psychological Review* 63: 81-97.

Mueller, G. E., and F. Schumann. 1894. "Experimentelle Beiträge zur Untersuchung des Gedächtnisses." *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* 6: 81-190, 257-339.

Nagasaki, H. 1987a. "Correlations of Stress and Timing in Periodic Tapping." *Human Movement Science* 6: 161-180.

Nagasaki, H. 1987b. "Frequency Dependence of Rhythm in Periodic Tapping." *Human Movement Science* 6: 247-256.

Nagasaki, H. 1990. "Rhythm and Variability of Timing in Periodic Tapping." *Human Movement Science* 9: 177-194.

Peters, M. 1977. "Simultaneous Performance of Two Motor Activities." *Neuropsychologia* 15: 461-465.

Povel, D.-J. 1981. "Internal Representation of Simple Temporal Patterns." *Journal of Experimental Psychology: Human Perception and Performance* 7: 3-18.

Povel, D.-J. 1984. "A Theoretical Framework for Rhythm Perception." *Psychological Research* 45: 315-337.

Povel, D.-J., and H. Okkerman. 1981. "Accents in Equitone Sequences." *Perception and Psychophysics* 30: 565-572.
Pressing, J. 1983. "Cognitive Isomorphisms between Pitch and Rhythm in World Musics." *Studies in Music* 17: 38.
Preusser, D., W. R. Garner, and R. L. Gottwald. 1970. "The Effect of Starting pattern on Descriptions of Perceived Temporal Patterns." *Psychonomic Science* 21: 219-220.
Restle, F. 1967. "Grammatical Analysis of the Prediction of Binary Events." *Journal of Verbal Learning and Verbal Behavior* 6: 17-25.
Restle, F. 1970. "Theory of Serial Pattern Learning: Structural Trees." *Psychological Review* 77: 481-495.
Restle, F. 1972. "Serial Patterns: The Role of Phrasing." *Journal of Experimental Psychology* 92: 385-390.
Restle, F., and E. Brown. 1970. "Organization of Serial Pattern Learning." *The Psychology of Learning and Motivation* 4: 249-331.
Rosenbaum, D. A., S. B. Kenny, and M. A. Derr. 1983. "Hierarchical Control of Rapid Movement Sequences." *Journal of Experimental Psychology: Human Perception and Performance* 9: 86-102.
Royer, F. L., and W. R. Garner. 1970. "Perceptual Organization of Nine Element Auditory Temporal Patterns." *Perception and Psychophysics* 7: 115-120.
Shaffer, L. H. 1982. "Rhythm and Timing in Skill." *Psychological Review* 89: 109-122.
Simon, H. A., and R. K. Sumner. 1968. "Pattern in Music." *Formal Representation of Human Judgment*, ed. B. Kleinmuntz (New York) 219-250.
Singer, A. 1974. "The Metrical Structure of Macedonian Dance." *Ethnomusicology* 18: 379-404.
Stetson, R. H. 1905. "A Motor Theory of Rhythm and Discrete Succession II." *Psychological Review* 12: 293-350.
Vos, P. G. 1977. "Temporal Duration Factors in the Perception of Auditory Rhythmic Patterns." *Scientific Aesthetics* 1: 183-199.
Vos, P. G., and S. Handel. 1987. "Playing Triplets: Facts and Preferences." *Action and Perception in Rhythm and Music*, ed. A. Gabrielsson (Stockholm) 35-47. Royal Swedish Academy of Music Publications 55.
Wickelgren, W. A. 1964. "Size of Rehearsal Group and Short-Term Memory." *Journal of Experimental Psychology* 68: 413-419.
Wickelgren, W. A. 1967. "Rehearsal Grouping and Hierarchical Organization of Serial Position Cues in Short-Term Memory." *Quarterly Journal of Experimental Psychology* 19: 97-102.
Wing, A. M. 1980. "The Long and Short of Timing in Response Sequences." *Tutorials in Motor Behavior*, ed. G. E. Stelmach and J. Requin (Amsterdam).
Wing, A. M., and A. B. Kristofferson. 1973a. "The Timing of Interresponse Intervals." *Perception and Psychophysics* 13: 455-460.
Wing, A. M., and A. B. Kristofferson. 1973b. "Response Delays and the Timing of Discrete Motor Responses." *Perception and Psychophysics* 14: 5-12.
Woodrow, H. 1909. *A Quantitative Study of Rhythm* (New York).
Woodrow, H. 1951. "Time Perception." *Handbook of Experimental Psychology*, ed. S. S. Stevens (New York) 1224.