

The Argument

This research is an instrumental investigation of a theory of rhythmical performance of poetry, originally propounded speculatively, in my *Perception-Oriented Theory of Metre* (1977). “Iambic pentameter” means that there is a verse unit consisting of an unstressed and a stressed syllable (in this order), and that the verse line consists of five such units. In the first 165 verse lines of *Paradise Lost*, there are two such lines.

Among other things, the theory takes up one of the central issues in metrical studies: all criteria for metricality hitherto proposed have been violated by the greatest masters of musicality in English poetry. The question arises, how do we recognize two verse lines that are very different in their structures as instances of the same abstract pattern of, e.g., iambic pentameter; and how do we distinguish a metrical from an unmetrical line. One great difference between this theory of metre and others concerns the status of deviation. Most theoreticians deploy a battery of tools to make deviant stress patterns conform with metric pattern. Only when all attempts fail, they speak of “tension”. When they succeed, they blur the distinction between e.g. Milton’s and Pope’s metrical styles. Or else, they have formulated different rules of metricality for Shakespeare and Milton, mainly on statistical evidence. But they don’t attempt to explain why some metric figure is acceptable for one poet but not for the other. By contrast, this theory actually welcomes deviances, and conceives of them as of essential parts of the aesthetic endeavour, just as metaphoric contradiction is an essential part of it. Linguistic stress pattern and metre may be conceived of as analogous to the two incompatible terms of a metaphor. The reader registers their incompatibility and resolves them in a pattern of performance. The utmost limit of rhythmicality (as of the meaningfulness of a metaphor) is the reader’s ability or willingness to cooperate, that is, to resolve the incompatibility of the two terms by a rhythmical performance (or, in the case of a metaphor, by a semantic interpretation). Thus, the approach to poetic rhythm advocated here is highly congruous with a wider aesthetics of “the elegant solution of a problem”, or of “the balance and reconciliation of opposite or discordant qualities”; and may account for poetic rhythm and metaphor by a homogeneous set of principles.

Thus, the perception-oriented theory of metre performs a small Copernican revolution, and instead in the verse structure, it places the constraints in the reader’s “rhythmic competence”: the utmost limit of rhythmicality is the reader’s ability or willingness to perform the verse line rhythmically. Such a formulation requires a systematic theory of “rhythmical performance”. The proposed theory of performance is based on Gestalt theory, speech research, and the hypothesis of limited channel capacity. Welles and Warren argue in their *Theory of Literature* (1956, Chapter 13) that in order to account for poetic rhythm, one must assume the existence of not one, but three metrical dimensions: prose rhythm, metric pattern, and performance

(generative metrists have reinvented the first two of them). It would appear that Wellek and Warren need the performance dimension in order to account for the fact that two unlike delivery instances may still be performances of the same metric structure, and to point out that some “sound-recorders” mistake in their analysis an accidental performance for the poem’s metre. For my purpose, performance is a perceptual solution to a perceptual problem, and as such determined by it to a considerable extent, but also leaving room for considerable creativity on the performer’s part.

The empirical study of the rhythmical performance of poetry must face almost insurmountable obstacles. “Rhythmical performance” is defined as the vocal conditions in which both the metric pattern and the linguistic stress pattern are simultaneously accessible to awareness. Since the former exists merely as a mental pattern, only the latter is available for an instrumental investigation. And even in this dimension, things are far from trivial. The main difficulty lies in the fact that in speech perception, contrary to common intuition, there is very little correspondence between what we *hear* and the shape of the sound wave as shown by the instruments. This discrepancy is not due to the machines’ incapability of representing the speech signal, but to the complex processing of the signal by the human brain. The sophisticated electronic instruments do give an accurate analysis of the sound information; but what really matters is its integration that takes place in the brain. Thus, for instance, the acoustic cues for linguistic stress are intonational inflection, pitch, duration, amplitude—in this order of decreasing effectiveness. It is impossible to predict from the machine’s output what their relative weight is, and which one of two consecutive stresses is perceived as stronger. This can be done only by the human ear.

The way out from this dead impasse was suggested by Polonius: “With windlances and with assays of bias / By indirections find directions out”. Empirical research must content itself with much simpler distinctions in the stream of speech, and reinterpret them in light of the theory propounded here, so as to indicate, indirectly, some mental biases and directions inaccessible to direct inspection. Gerry Knowles of Lancaster University created, independently from my plight, precisely the tools which I needed. In his 1991 paper, he investigated the nature of tone-group boundaries. He distinguished internally defined prosodic patterns and external discontinuities at the tone-group boundaries. The former consist in some consistent f_0 pattern used in ordinary speech; the latter are temporal discontinuation (pause), pitch discontinuation (a sudden change in f_0) and segmental discontinuation (that is, in normal speech the articulation of adjacent words is overlapping; when there is no overlap, it may be perceived as discontinuity, even if there is no pause). This would be the most evasive type of discontinuity. “The important distinction that seems to be emerging is between boundaries with or without pauses”. As shown in Chapter 3, these distinctions may account, in a quite straightforward manner, for the rhythmical performance of enjambment; but also, less directly, for the rhythmical performance

of strings of stressed syllables, and still less directly, for that of a stress maximum in a weak position.

In another paper, Knowles (1992) explores the alignment of the f_0 contour with vowels and consonants. “Although the effect of a tone might be to highlight a whole word or phrase, its focus is on a single syllable. Within the syllable it focuses on the vowel, and if the vowel is a diphthong, on one of the elements of the diphthong. Ultimately within the relevant vowel there is a single point which appears to be the focus of accentuation” (Knowles, 1992: 294). Such points may be located in various places in the vowel. Accordingly, he speaks of early peaking and late peaking, as the case may be. In our investigations, late peaking turned out to be the source of an impetuous forward drive in the perceptual dynamics of the verse line, and a major resource for rhythmic grouping. This research explores how the correlates of stress and of tone-group boundaries can be exploited as conflicting cues for the perceptual accommodation of the conflicting patterns of speech and versification.

In the verse instances discussed in the present study, the stress pattern considerably diverges from the metric pattern. In such cases, the reader of poetry must rely on his metrical set. That is, whenever metric regularity is suspended, the reader may *echo*, so to speak, in his short-term memory, the regularly alternating underlying beats, even though they may have no trace in the acoustic signal. The reader may compensate, to some extent, for the absence of the metrical signal, by *anticipating* the return of regular beats. All this is possible, if at all, for a very short period only, over a span of a very few “chunks”. I said *if at all*, because the extra mental space required may not be available; besides, it takes a fairly experienced reader to perform such verse rhythmically.

The present approach assumes that poetic rhythm is, essentially, an auditory phenomenon, though also affected by syntax and semantics. Its auditory qualities may be accounted for if we assume that it is processed in short-term memory, which functions in the acoustic mode, and is constrained by its limitations. The contents span of short-term memory is limited to seven monosyllabic words plus or minus two (Miller, 1970; that is why the longest verse line that can be perceived as a rhythmic unit without an obligatory break is ten-syllable-long). Its time span is roughly the period we can remember, e.g., a telephone number without rehearsal. During this period short-term memory functions like an echo box. In order to render a verse line perceptible as a rhythmic whole, the reciter must manipulate his vocal resources in such a way that the verse line can be completed before its beginning fades out in short-term memory. When the immediately observable string of syllables deviates from metric regularity, the metric pattern may be perceived as reverberating in the background, provided that sufficient mental processing space is available. Training cannot expand these spans of short-term memory. The only thing one can do is to recode the verbal material in such a way that it occupies less mental processing space. Thus, for instance, “a man who sells goods” can be recoded as “merchant”, and “a merchant who sells meat” can be recoded as “butcher”. Such re-

coding is impossible in poetry, where the actual words may not be changed. Still, some mental processing space may be saved by two kinds of vocal manipulations: grouping and clear-cut articulation. Gestalt theory has laid down fairly rigorous rules of what facilitates perception: these include grouping and parsing (which is one kind of articulation). Accordingly, we may expect reciters to over-articulate, on the one hand, word and syllable boundaries (parsing) and, on the other hand, to group syllables and words in certain ways. Speech research of the past thirty years (cf. Lieberman, 1967) has established that in the flow of everyday speech we tend to rather careless articulation, and the listener has to do in the course of decoding a lot of subliminal guesswork. In conversational speech words are normally run one into the other in English (and more so in French), and it takes special decoding effort to determine the word endings. Taylor (1990: 212) suggests: "Say rapidly 'How to wreck a nice beach', and it will sound like 'How to recognize speech' [...]. The sentence illustrates the point that word boundaries are anything but fixed. Without boundaries words are hard to recognize". Thus, much decoding effort can be saved by clear articulation of word endings. Clear-cut articulation of phonemes and of syllable and word boundaries may save a lot of mental processing space. Intonation is a typical means of over-articulating syllable and word boundaries; but the over-articulation of the syllable (or word) final consonants too contributes to the over-articulation of boundaries.

There are less and more marked instances of deviance. Two consecutive stresses in the iambic metre must be considered as deviation. When a compound like "black-bird" (whose first syllable is more strongly stressed) begins in a strong position, it is less marked than when it begins in a weak position. In Pope's poetry, all such compounds begin in strong position (except in the first position of the line, as in "Long-sounding islands"). In Book I. of Milton's *Paradise Lost*, 15 out of 20 such compounds begin in a weak position; there is a similar ratio in Shakespeare's Sonnets. In Shelley and Keats, slightly more than half of such compounds begin in a weak position. In such phrases as "black bird" where the last syllable bears the strongest stress, the unmarked form is when they end in a strong position. Some metrists even rule instances which end in a weak position as unmetrical. In Pope, all such strings end in a strong position (I found only one exception, "Awake my *Saint John*", which Pope probably pronounced as "Sinjen"). In Shakespeare, Milton, Shelley, Keats there are some such strings which do end in a weak position). When a string of stressed syllables ends in a strong position, a construction like Donne's "Shall behold God" (where the stressed syllable in a weak position is part of a polysyllabic which has an unstressed syllable in a weak position) is more marked than "Shall be old God" would be. Indeed, some metrists rule such instances as unmetrical. The latter, in turn, is more marked than "Shall see old God", where the stressed syllable in a weak position is enclosed between two stressed syllables in strong positions, and the stress pattern of natural speech conforms with the iambic pattern. There are, then, scales of markedness, on which each poet draws his own utmost limit of metricality. The more marked a deviation, the more emphatically

the devices of grouping and over-articulation will be deployed. But the present research found plenty of evidence that leading British actors tend to have recourse to such devices even in less marked instances, where less marked performance patterns would be available as well.

Halle and Keyser say that an unmetrical line is a line in which a stress maximum occurs in a weak position (a stress maximum is a stressed syllable between two unstressed ones: “a garden” contains a stress maximum; “a big garden” does not). Halle and Keyser and their critics found eleven instances of unmetrical lines under this theory in major English poetry. My 1977 book added a list of over forty more instances. Two thirds of the instances occur in the seventh out of four positions available for “violation” in an iambic pentameter line. In Milton’s “Burnt after them to the bottomless pit”, *-bot-* is a stress maximum in the seventh position. The present theory predicts that in such instances the last four syllables (“bottomless pit”) will be emphatically grouped together, foregrounding a closed, symmetrical shape of two stressed syllables enclosing two unstressed ones (“stress valley”). Such a performance requires segregation of the stress valley from the preceding portion of the line; at the same time, continuity of the phrase must be preserved. This can be accomplished by having recourse to conflicting cues. When such a stress valley begins in the seventh position, it threatens the integrity of the line; regularity is reinstated precisely on the last strong position of the line, generating a powerful closure. In terms of our foregoing discussion, the computer cannot show a stress valley; but it can show conflicting cues for continuity and discontinuity where syntax would require an uninterrupted stream of speech.

A distinction must be made between performance patterns and their acoustic and phonetic correlates. The relationship between, e.g., a stress valley and its acoustic correlates is similar to the relationship between a phoneme and its acoustic correlates. We are interested in the phoneme as an abstract category, and ignore the specific acoustic cues that are its exponents. Consequently, there is usually a trade-off between the possible acoustic correlates that may cue a certain phoneme. Thus, for instance, a voiced stop may be cued by the straightforward activation of the vocal folds, or by a lengthening of the preceding sonorant, or by reducing voice-onset-time, or by aspiration. Most language users would not distinguish between the various vocal devices; they merely perceive a unitary abstract category, such as [b, d, g]. The same is true of the perceptual organizations required by the rhythmical performance of a deviant verse line. Consider the case of a stress valley, produced to accommodate a stress maximum in a weak position. Since a stress maximum occurs, by definition, in mid-phrase, or even in mid-word, the performer will face the following conflicting tasks: he must segregate the stress valley from the preceding context, but must preserve the continuity of the phrase, or even more so, of the word. The listener, or even the reciter himself, will be aware at best that the perceptual problem has been solved. On closer inspection, they might discern that a stress valley has been applied. On still closer inspection, they might even discern the opposite tendencies of continuity and discontinuity between the stress valley and the rest

of the phrase or word. But it is impossible for them to discern by what phonetic means this has been accomplished. And, in fact, what matters for the solution is the abstract category “stress valley”, and it is immaterial what trade-off between the various acoustic cues may take place.

The acoustic cue for continuity is usually quite straightforward: there is no measurable pause before the stress valley. Quite frequently also “an internally defined intonation pattern” is assigned to the sequence of four syllables. It is more difficult to discover the cues for discontinuity. I contend that there is an open list of possible acoustic cues. Reciters display an astonishing degree of creativity; new performances provide acoustic cues some of which are quite expected, but some are entirely unforeseen. But as long as they generate a unitary perceptual category and indicate the required segregation, listeners exposed to them for the first time immediately recognize them as appropriate (assuming that continuity is taken care of).

We have no access to what happens in that black box, the reader’s head; we have access only to vocal performances. We can only make inferences from these performances to mental processes. And the vocal performances reflect the constraints of three kinds of competences, each later one relying on the preceding one: the competence to identify the conflicts between stress pattern and metre; the competence to find a solution to the conflict, and the proper command of voice to carry out the solution. When the performance of a deviant verse line is judged rhythmical, we may assume that the reciter had command of all three competences; when not, we may make only more or less accurate guesses as for which one(s) of the competences failed.

This work assumes, then, that when stress pattern and metre conflict in poetic rhythm, the reader may accommodate them in a rhythmical performance. Rather than in verse structure, the constraint for acceptability in versification is placed in the reader’s “rhythmic competence”, his ability or willingness to perform the verse line rhythmically, that is, in a way that both metric pattern and the linguistic stress pattern should be accessible to awareness at one and the same time. It suggests a cognitive mechanism that may render this feasible, and points out the principles of the vocal manipulations required. If the reader succeeds in such a performance, increased tension is perceived; if not—the verse line disintegrates and tension ceases. Halle and Keyser’s notion of stress maximum in a weak position is a powerful tool to describe a very high degree of deviance in versification; but it does not necessarily render a verse line unmetrical. The existence of a corpus of about 60 verse lines with a stress maximum in a weak position in major English poetry suggests that such verse lines may be acceptable on *some* grounds. The fact that about two thirds of the instances occur in precisely the seventh out of four positions available for “violation” suggests that these deviances are not random. The approach advocated here presumes to explain the cognitive rationale of this distribution. Two further points are demonstrated: that experienced readers tend to agree upon the kind of performance demanded by such verse lines; and that these performances are in harmony with the

expectations of the theory propounded here. Moreover, when alternative mappings of stress pattern to metric pattern are possible in a verse line, some highly experienced readers may prefer a mapping that involves a stress maximum in the seventh position to some other, perfectly “metrical” mapping under the stress-maximum theory. My detailed discussion of how experienced readers handle instances of stress maxima in weak positions is the test case of the validity of my challenge to the existing paradigm.

*

The first three chapters of this book give the general theoretical framework of this study. Chapter 1 presents the essentials of my original perception-oriented theory of metre, indicating, in a synchronic perspective, the theoretical problems which it set out to solve, as well as placing differences between metrical styles in a diachronic perspective. Chapter 2 bestows an operationally definable psychological meaning upon the statement “This is a rhythmical performance of this verse line”. It draws upon Gestalt theory, speech research, and adopts the limited-channel-capacity hypothesis. The key word in its various sections is “simplicity”. Chapter 3 recapitulates the notion “rhythmical performance” in an empirical perspective, and makes a further distinction between delivery styles, indicating the deployment of different articulatory strategies entailed by them. It is argued that poetic rhythm is inaccessible to empirical research, and one can make only roundabout inferences about it. Empirical research can be applied to what appear to be rather trivial elements; but these elements may gain significance in light of the theories expounded in the preceding two chapters. In this way, distinctions that are tailor-made for the enjambment could be applied, less straightforwardly, to strings of stressed syllables and to stress maxima in weak positions as well. This chapter also discusses the Gestalt notion of “perceptual forces” on the subphonemic, the syntactic and the vesification-unit levels. The latter two levels would more properly belong in Chapter 7, but the first level (“early and late peaking”) clearly belongs here, and the syntactic level too has an empirical ingredient; so it is more parsimonious to discuss the three levels here, together.

Each one of Chapters 4-8 is devoted to the rhythmical solution of performance problems arising from one kind of complexity or structural feature. Chapter 4 scrutinizes the definitions of caesura, and explores its effects on the perceptual dynamics of the line. Then it observes in detail how leading British actors handle the issues involved. Chapter 5 discusses at great length the theoretical and perceptual problems concerning a variety of metric configurations, in each of which there is a string of at least two consecutive stresses. It discusses the arising problems of performance and attempted solutions in a wide range of delivery instances. There is also a brief discussion of some theoretical problems concerning the demarcation line between

“metrical” and “unmetrical”. Chapter 6 is, in an important sense, the touchstone for my challenging of the prevalent paradigm in metrical studies. Halle and Keyser offer “stress maximum in a weak position” as the criterion for an unmetrical line. It is pointed out in this chapter that there are at least 60 instances of stress maxima in weak positions in major English poetry, and that their distribution in the verse line is far from being random. It is suggested that this distribution is in harmony with the predictions of the present theory, and further predictions are made as for the performance patterns to be offered to pentameter lines with a stress maximum in the seventh position by competent readers. It is claimed that the performances of competent readers of such lines are, indeed, in harmony with these predictions. What is more, some very competent readers tend to have recourse to such solutions even in verse lines in which other solutions, perfectly metrical under the Halle-Keyser Theory, would be available as well. The discussion is preceded by an appraisal of the cognitive parsimony of the Halle-Keyser theory.

Chapter 7 is devoted to enjambment. Rather than reviewing all the theoretical problems involved, this chapter takes its departure from the controversy whether conflicting intonation contours can be indicated in vocal performance. The present study takes sides with those who answer this question in the positive, and supports this judgment by readings of experienced readers. These readers do, indeed, have recourse to conflicting cues to indicate discontinuation at the line ending and, at the same time, continuation of the run-on syntactic unit. Chapter 8 faces the most disconcerting issue in our present inquiry, bisyllabic occupancy of metrical position or, in plain English, instances in which two syllables must be assigned to one metrical position. Halle and Keyser laid down the phonetic conditions in which this can be done. English poets, even Donne in his *Satyres*, seem to obey the Halle-Keyser rules in this respect. The present work views the issue as a perceptual rather than a phonetic problem. The conditions laid down by Halle and Keyser facilitate the pronunciation of the two syllables in such a way that the listener hears two syllables, but is aware that there is only one metrical position underlying them. So far we have found that the over-articulation of word boundaries was one of the most effective means for allowing the listener to perceive the one-to-one correspondence of the regularly alternating weak and strong positions and the irregularly alternating unstressed and stressed syllables; when this one-to-one correspondence must be avoided, the rhythmical performance requires precisely the opposite, that is, under-articulation of the boundary between the two syllables assigned to one position.

Chapter 9 explores the ways in which music can throw light on poetic rhythm. There is a long tradition of metrists who assume that poetic rhythm is founded on equal or proportional time periods; some of them even use musical notation for rhythm. It is argued that all measurements contradict this assumption; even where measured time *is* significant, it does not signify such overall qualities as equal or proportional time periods, but rather such local qualities as stress or discontinuation. The present study subscribes to the rival research tradition, which assumes that poetic rhythm is based on an abstract mental pattern; as long as the abstract pattern is

available to awareness, irregularities in the immediately observable linguistic units can be tolerated. While in music pause replaces a note, in poetic rhythm, with a few exceptions, pause is perceived as an event intruding upon the verse line as a whole, and forces the line to reassert itself in the listener's perception. Even very long pauses are tolerated, if the integrity of the verse line is reinforced by a variety of vocal manipulations. Paradoxically enough, it has been found that it is rather the diatonic aspect of music that may have a significant contribution to the perception of poetic rhythm: in some instances, cadential intervals between syllables may enhance the perception of metric boundaries. The short excursus on Hungarian poetry in Chapter 10 demonstrates that the same strategies found in the performance of English poetry are used, in a different theatrical tradition, to solve considerably different problems that arise in Hungarian poetic rhythm. The Appendix places the present study in a much wider methodological perspective. On the one hand, it raises the problem of reductionism in interdisciplinary studies: it argues that problems in poetics cannot be reduced to the "more basic" sciences. On the other, it faces the problem how the perceived affects of poetry can systematically be related to poetic structures. It widens the scope of inquiry beyond metrics, to other issues concerning the sound effects of poetry, such as the emotional symbolism of sound patterns. It suggests that going out from poetics to phonetics or acoustics may, more often than not, merely multiply information and obscure rather than explain issues, and attempts to define the conditions in which reliance on these "more basic" sciences *can* account for the perceived effects of poetry. Empirical studies of poetic rhythm that make use of instrumental phonetics must meet these methodological conditions.