

A Conversation with Reuven Tsur on Intonation and Prosody

Natalie Gerber and David Nowell Smith: The conceptual vocabulary of prosody as it has been passed down over centuries, more or less intact, is a rather blunt instrument given the intricacy of the phenomena it is charged with analysing. How in your own work have you managed to develop alternative conceptual schemes and terminologies in order to grasp rhythmical performance in all its complexity?

Reuven Tsur: The term ictus refers to an indefinite mixture of linguistic stress and metrical downbeat. In traditional metrics ictus is marked as – over the syllable crest (\bar{a}), non-ictus as u (\tilde{a}); iambic pentameter is marked as ‘Āt ōnce ās fār ās Āngĕls kĕn hĕ vĭĕws’. This method has two setbacks. First, students are confused—and you can’t blame them for that—when they have to decide, in assigning prosodic marks, whether to mark a linguistic stress or metrical downbeat as ictus. When in the iambic metre, for instance, an unstressed syllable happens to be even-numbered in its sequence, they have no rule to decide whether to mark an ictus or non-ictus. The result is utter confusion. Second, and more serious, this is an efficient way to regularise metre but, by the same token, also to efface any complexity in poetic rhythm.

So when I started out in my inquiry into poetic rhythm, the first thing I did was to clearly separate the two kinds of strong events. I used the above two symbols to mark only metric downbeats and upbeats, and wrote them under the text. I marked

linguistic stress, over the text, with the mark /. I assigned the sequence of metric upbeats and downbeats as regularly alternating, irrespective of the kind of syllable they referred to. And I marked linguistic stress irrespective of whether it corresponded to a metric downbeat or upbeat. At the time I couldn't formulate it so explicitly, but that's how I acted intuitively. The explicit formulation came to me when I was exposed to the Halle-Keyser theory. I realised that they had devised a system very similar to mine, but their formulations and definitions were incomparably clearer and more unambiguous. I also adopted the terms metrical weak and strong positions for metric upbeats and downbeats, and also their marking of w and s positions (for weak and strong) under the text, instead of the traditional symbols.

In spite of our identical apparatus, Halle and Keyser's purpose was very different from mine. In the spirit of generative linguistics, they set out to discover the rules that can generate all metrical lines and no unmetrical ones. This involved them in a circular, irrefutable argument, in Popper's sense.

NG & DNS: You mean: because any counterexample is simply deemed to be 'unmetrical', and hence not a counterexample at all?

Exactly!

How can we know whether a line is metrical or not? In practice, the answer boils down to the following. A metrical line is one that conforms with the rules of the theory. Thus, a counterexample cannot refute the theory, because it is unmetrical by definition. This part of the theory did not greatly disturb traditional metrists, because most of them, too, were interested in which deviations were allowable, and which not.

In my approach, by contrast, the new marking system allowed me to distinguish between a convergent style (where linguistic and metrical downbeats largely coincide), whose prototypical representative is Alexander Pope, and a divergent style, whose prototypical representatives are Milton and Shelley. My method visually foregrounded the simultaneous perception of the patterns of language and versification, where they coincide and where they conflict, assisting my attempts to explore the nature of a rhythmical performance. Halle and Keyser, by contrast,

explicitly state that they have nothing to say on how the deviations acceptable by their theory should be performed.

NG & DNS: Presumably this is not just a question of a poet's 'style', but also of their position within a tradition: these patterns are historically embedded in important ways.

This 'position with a tradition' has very good reasons. When at the early twentieth century recording machines were introduced, many researchers, fascinated with the new possibilities, measured relationships in recorded readings and thought they were measuring relationships in the poems, whereas, in fact, they were measuring relationships in accidental performances. Wellek and Warren and Wimsatt and Beardsley attacked this approach systematically, but it still persists in the form of 'proportional' or 'equal timing', in spite of accumulating evidence to the contrary (in *The New Princeton Encyclopedia of Poetry and Poetics* (1993) it is still presented as a valid approach). I have proposed to avoid this fallacy not by banning performance from discussion, but by treating it as a problem-solving activity, constrained by the linguistic and versification patterns, on the one hand, and the reader's and listener's cognitive processes, on the other.

NG & DNS: So your main argument with Halle and Keyser is also that with a much longer tradition which disregards 'descriptivism' and 'performance'?

I think we should emphasise that, in spite of all this, the Halle-Keyser theory has enormous merits. It's easier to communicate and even think about metrics after Halle and Keyser than before. Their magic key to tell apart 'metrical' from 'unmetrical' lines is their theoretical construct 'stress maximum'. This is a lexically stressed syllable between two unstressed ones, within the same verse line and the same syntactic unit. 'A gárden' contains a stress maximum, 'a róck gárden' does not. According to Halle and Keyser (1971), all deviations from metric regularity are acceptable, except a stress maximum in a weak metrical position, which renders a verse line 'unmetrical'.

To understand how the Halle-Keyser theory facilitates thinking about versification, I shall invoke George Miller on recoding information in order to save mental processing space; one way to do this is 'to group the input events, apply a

new name to the group, and then remember the new name rather than the original input events'. The stress-maximum theory has stimulated numberless new insights in prosodic research all over the world, in spite of the fact that it merely coded all the conditions for the acceptance of stressed syllables in weak positions (which had been well known to prosodists for centuries) into one well-articulated, easily manipulable concept. This is a clear-cut case of 'grouping the input events, applying a new name to the group, and then remembering the new name rather than the original input events'. The term 'stress maximum', its definition and conditions of application are stored in long-term memory; but when the term gets into working memory, it is, as I said, easily manipulable, and retrieves, with great speed, only the relevant input events from long-term memory, thus requiring relatively little mental working space. When confronted with different instances of metric deviation, it retrieves different conditions — thus involving rapid 'aspect-switching'. That is why it is so conducive to insights. I certainly could not have had some of my own exciting insights without the help of this term, even though I object to Halle and Keyser's basic agenda. To give just one example. If I hadn't known that a stress maximum in a weak position is said to render a verse line unmetrical, I would not have discovered that some strong positions are stronger than others, and could not have created a scale of mismatches, and many other things.

Halle and Keyser's theory posed a threat to many traditional prosodists: it robbed from them their privilege to say anything and be right. There was a concerted effort all over the world to find counterexamples to the theory. Together with Halle and Keyser's own counterexamples something like twelve were found in major English poetry. That's not very bad. However, during the years I found about 55 counterexamples. This finding does much more than increase the numerical weight of the counterexamples: it allowed me to discover that there is some logic in the distribution of violations over the line. There are four positions in an iambic pentameter line available for violation by stress maxima: the third, the fifth, the seventh and the ninth. A random distribution would allocate roughly twenty five percent of the violations in each of them. About two thirds, however, occur in position seven, nearly one third in position three; I have found only a few, sometimes doubtful, examples in positions five and nine. I will later try to account

for this distribution. In the meantime, I came to the conclusion that perhaps all deviations are acceptable, provided they can be performed rhythmically, and that the boundary of acceptability should be transferred from verse structure to the performer's rhythmic competence: his ability and/or willingness to perform the deviant structure rhythmically. More important, I came to the conclusion that for some reason a stress maximum is more easily performed rhythmically in the seventh than in the fifth or ninth position.

NG & DNS: So it sounds as though your engagement with Halle and Keyser really drove your sense that a cognitive explanation was required for rhythmical phenomena?

On the contrary. I came to that conclusion long before being exposed to Halle and Keyser. Later, when I was tackling with the Halle-Keyser theory in Hove (Sussex), a friend active in the Poetry Society called from London, and said that there is here a chap who says he has a metrical theory, and will give a lecture that night at the poetry society. When he told me that the chap was Jay Keyser, I took the first train and went to London. We had a fascinating encounter that had a crucial effect (not through agreement but disagreement) on the future development of my theory. His most important comment was that if I argue that all deviations are acceptable, provided they can be performed rhythmically, I must come up with a systematic theory of performance. Otherwise, my proposal is empty (he wasn't very choosy in his terms).

My first reaction to this proposal was shock. But later I realised that he had been right. I started with introspection. From my early teenage years, I had been addicted to poetry recital. After WWII there was famine in Budapest, and many people fled to the country towns to eat. To our town came a couple of leading Jewish actors, Holocaust survivors back from the camps, who revolutionised our local theatre. However, after a short time they found themselves out of a job for political reasons. My mother, president of the local Jewish women's committee, helped them in many ways: among others, she sent me to the wife to have lessons in poetry recital. I learnt from her a lot, from wide generalizations to trivial exercises. One of the most trivial ones was to close my teeth and recite a poem, another one was to overarticulate each consonant in reciting a poem. These exercises were meant to enhance the

articulation of vowels and consonants. This was a piece of practical wisdom, without far-reaching aesthetic or psychological considerations. Many decades later, however, I understood that clear-cut articulation also saved mental processing space, to make room for the simultaneous processing of conflicting patterns. Eventually I came to the conclusion that I had to master two new disciplines from scratch, Gestalt Psychology and the results of instrumental phonetics (in the pre-personal-computer era I wouldn't have dreamt of conducting the experiments myself).

NG & DNS: Why these two disciplines in particular?

When I wrote in the early seventies to D. B. Fry of UCL, seeking his help in my attempt to obtain empirical support for an instrumental analysis for my fledgling theory of rhythmical performance of poetry, he answered:

It is clear to me that the ear and the brain are the only satisfactory instruments for dealing with matters of rhythm and metre. Not because their effects are any more 'subjective' than any others, but because they represent a complex of impressions for which ear and brain are the only appropriate transducing instruments. So in fact I think you are better off listening to and thinking about metre than in doing physical measurements. [...] All this is not to say that one cannot learn something from physical measurements, but in the case of metre this is very much a matter of following one's hunches'. (personal communication, 22 November 1972)

Nevertheless, he offered me his help; but when I sent him my stuff, I got no answer. A few weeks later all the linguistic journals were printed in black frames, announcing his death. It took me twenty-five years to learn to ask questions that the machine understands.

NG & DNS: How were you finally able to do so?

My first direct exposure to instruments happened when I was on a sabbatical at Yale. At a Hanukka party I met Al Liberman, president of the Haskins Laboratories for speech perception, whose works I knew from books and journals. He was surprised that a literary person was interested in their work (later I found out that professors of comparative literature in Yale had no idea about the very existence of

the Laboratories, five minutes walking distance from their department). He invited me to try my theories on their equipment, and asked Terry Halwes to initiate me into the mysteries of speech research. I learnt an enormous lot there; among other things, it resulted in my book *What Makes Sound Patterns Expressive?—The Poetic Mode of Speech Perception*, published in the Roman Jakobson Series of Duke University Press. But in the end, regarding the rhythmical performance of poetry, I realised that—alas!—Fry might be right.

The problem through all these years was that there is an enormous discrepancy between what we experience when listening to speech and what can be read off a spectrogram. In fact, as Fry implied, 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.

NG & DNS: So perception is cognitive through and through?

This question one must answer in two stages: speech perception and rhythm perception. According to the 'motor theory of speech perception', you will wonder, speech perception is pre-cognitive. When we utter a word, we make certain articulatory gestures. We transmit a stream of sound information, not about the speech sounds, but about the articulatory gestures, which the listener's articulatory system decodes as the intended speech sounds. Rhythm processing is cognitive. That is why the shape of the machine's output resembles neither the speech sounds', nor the rhythm's shape.

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. Since then, every few years I applied to universities that had advanced phonetic laboratories for help in my research. Usually I received the answer that they had no staff experienced in this kind of research. Eventually I ended up at the Department of Linguistics and Modern Languages, Lancaster University; there I was exposed to British phonetician Gerry Knowles's work, who quite independently from my plight had created precisely the tools which I needed. The 'state of the art' has not greatly changed since that time; it is rather that I found a way to ask the machine questions which it can understand, and the answers to which make sense within the

framework of my theory.

At the time of my encounter with Jay Keyser I was a research fellow at Southampton University, comparing seventeenth-century English Metaphysical poetry with eleventh-century Hebrew poetry in Spain. In my recreation time, instead of playing tennis, I tried to work out my intuitions regarding English tonic-syllabic metre. The morning after my meeting with Jay Keyser, I ran into the head of the English Department, the great traditional Milton scholar, T. F. Prince. Still fascinated with our encounter, I told him of the stress-maxima theory. To make sure he had understood, he quoted from *Paradise Regained* the line: 'And made him bow to the gods of his wives'. Three stress maxima occur in this line; two occur in strong positions, but we are interested here in the third one: 'the gods of'—a noun that bears lexical stress, between two function words, that bear no lexical stress; and it occurs in the seventh, weak position. When I went home, I tried to find the exact reference for the line, but had no concordance at home. So I re-read the whole *Paradise Regained*, with an eye on stress maxima in weak positions. To my pleasant surprise, I found no fewer than eighteen 'unmetrical' lines under the stress-maxima rule in this single work (as against the twelve examples found by a concerted effort in all English poetry). Now consider this. This huge number of 'violations' does not render Milton an incompetent poet; on the contrary, he is deemed by many as one of the most musical English poets. Within my corpus, Milton had the greatest number of stress maxima in weak position, and had the greatest number of metric deviations in general (except Shelley), and precisely these two poets are usually counted among the most musical poets in the English language.

NG & DNS: What do you mean by 'musical' here? It seems quite a problematic term, given that 'music' itself includes (just in the Western classical tradition) monody, Mozart and Messiaen. And throughout history readers have complained of Milton's 'lack of music'. More broadly, is there a danger in appealing to intuitions (this 'feels' musical) without looking at how these intuitions are formed?

This is a rather perplexing question. I know that many people speak of musicality in verse; they know what they are talking about, but cannot explain what it is. Strictly speaking, it is a metaphor. I suppose one ingredient is 'rhythmical' in the sense

discussed here. Another ingredient concerns the fusion of precategorical elements, phonetic and semantic. I have discussed these issues at exceptionally great length in my book (2012) *Playing by Ear and the Tip of the Tongue: Precategorical information in poetry* (Amsterdam / Philadelphia: John Benjamins).

There is experimental evidence that some people tend to hear, in certain conditions, the acoustic features of speech sounds reverberating in short-term memory, whereas some tend to hear them as compact, 'opaque' speech sounds. And some conditions are typically more conducive to reverberation than others. Musicality is associated with reverberation. Let's consider two aspects of conditions conducive to the perception of words as compact or reverberating: lateral inhibition, and clear-cut vs. blurred boundaries.

When two sound patterns are very similar, they enhance in each other the reverberation of the rich, precategorical information (acoustic information, phonetic features). This is part of the meaning of 'musical'. When two sound patterns are slightly similar, they inhibit reverberation in each other; speech sounds are perceived, rather, as compact. A chapter in one of my books is called 'Musicality in Verse, and Phonological Universals'. I took the first part of the title from Kenneth Burke, the second part from Roman Jakobson. It took me years to discover that the title contained the stem 'vers' twice. If, however, you take care to pronounce this syllable both times at the same pitch and length, you suddenly *hear* the compelling alliteration. 'Lateral inhibition' is a necessity of efficient vocal communication, and has well-documented experimental literature. It prevents sound from directing attention away from the meaning to be conveyed.

According to one of the Gestalt rules of perception, interaction between gestalt-free elements (colour in visual perception and overtones in music) is increased within, but inhibited across, strong gestalt boundaries; but it is boosted across weak Gestalt boundaries. In visual perception, Ehrenzweig demonstrates this by colour interaction. In aural perception, he demonstrates it in 'true polyphony' or in a fugue, where, if overlapping strong melodic shapes blur each other, the overtones fuse, generating an undifferentiated mass of sounds.

Consequently, in Pope's convergent poetry and Milton's divergent poetry alliteration assumes a different character. Consider the following two excerpts, from

Edgar Allen Poe and Milton:

Chilling and killing my Annabel Lee.

Poe: 'Annabel Lee'

Of Man's first disobedience, and the fruit
Of that forbidden tree whose mortal taste
Brought death into the World, and all our woe,
With loss of Eden, till one greater Man
Restore us, and regain the blissful seat,
Sing, Heavenly Muse, [...]

Milton: *Paradise Lost*

Where is alliteration more intensive? More likely than not, we will get the answer that 'chilling and killing' is more intensive. But consider this. In Milton's passage there is an unusually great number of intensive alliteration patterns. F-r-t in *first* and *fruit*; s-t in *first*, *taste*, *restore*, *seat*; l-s in *loss* and *bliss*; *Eden* is wholly included in *-bidden* which, in turn, is wholly included in *-bedience*. Now notice this too. In the Poe excerpt, stressed syllables occur in all and only strong positions, unstressed syllables in all and only weak positions. The alliteration pattern consists of two symmetrical whole words that comprise identical speech sounds (all but one) in the same order. The alliteration patterns reinforce metric regularity. In Milton's excerpt, by contrast, stressed syllables occur in mainly, but not all and not only strong positions; unstressed syllables occur in mainly, but not all and not only weak positions. Consequently, Poe's metric shape is convergent, much stronger than Milton's. Conversely, Milton's verse is highly divergent, has much weaker metric gestalts. What is more, Milton's passage displays an exceptionally complex enjambment. It begins with a preposition (of) predicting a verb, which occurs six lines later (Sing). All the time, the reader must remember he has a loose end, while following the proceeding of the text. In addition, we have such minor enjambments, as 'and the fruit / Of', or 'mortal taste / Brought', or 'till one greater Man / Restore us'. Consequently, the Milton passage has exceptionally weak gestalts on all levels. This boosts the interaction of the pre-categorical phonetic and acoustic features of the alliteration patterns across all those boundaries. Briefly, in Poe's text, alliteration is

focused on two compact words, perceived as playful punning rather than diffuse musical texture. Milton's alliterations, on the contrary, are backgrounded as rich, diffuse, but fused, pre-categorical, 'musical' texture.

To come back to the question of musicality: if Tennyson speaks of Milton as the 'God-gifted organ-voice of England', whereas some other people 'have complained of Milton's 'lack of music', it does not mean that in matters of musicality everybody is entitled to his or her arbitrary intuitions, nor that one is right and the other wrong. One can give a structural description to the different intuitions. It would suggest, according to the present conception, that they draw at different points on the scales of complexity the utmost limit of their ability or willingness to do all the required cognitive processing. If performed in a certain way, it's 'God-gifted organ-voice'; if not, it 'lacks music'.

NG & DNS: So musicality can't simply be mapped onto either convergent or divergent styles?

In due course I had to make an additional distinction between deviations that enhance musicality, and deviations that violate it. Donne in his Satires, but not necessarily in his Holy Sonnets, has an overwhelmingly greater number of metric deviations in general and stress maxima in weak position in particular than either Milton or Shelley. Nevertheless, Donne is not deemed as a musical poet; on the contrary, he is deemed as exceptionally unmusical, and perhaps purposefully so. The numerical difference between Pope (my prototypical convergent poet) and Milton (my prototypical divergent poet) was 24 mismatches per 100 lines (that is, 1000 positions). It is not only the number, but also the relative markedness of mismatches that make the great perceptual difference. Shelley had a somewhat greater proportion of mismatches than Milton. In Donne's Satire II there were 100 (!) more mismatches per 100 lines than in Shelley's 'Adonais'.

Eventually, I had to come to the conclusion that the greater the number and severity of deviations, the more musical the poem, provided that it is performed in a certain way, and up to only a certain theoretically-undefined point; beyond that point, the deviations escape the performer's control, and render the poem exceptionally unrhythmical. Later, as I indicated above, I found that not all the weak

positions are equally available for violation. In some there is an abundance of stress maxima, and in some only a few doubtful cases. This would suggest that in some weak positions it's easier to perform a stress maximum, and in some — more difficult. Thus, a scale of mounting difficulties can be constructed, implying, perhaps, a scale of mounting tension.

NG & DNS: So you see the musicality of verse not just in terms of deviation/variation, but also the aesthetic pleasures involved in the problem-solving constraints of performance?

Yes, indeed.

This is done by recoding. Based on the foregoing distinctions, Gestalt Theory and the instrumental analysis of speech, I tried to propose what are the 'certain ways' of performing rhythmically a line that contains a stress maximum in a weak position. But before doing this, I must make a few more comments on limited channel capacity, and briefly map out the metric boundaries in a verse line.

Cognitive psychologists assume that the cognitive system has limited channel capacity. In order to increase the number of streams of information processed in a channel at any given time, one must reduce the load on the other streams. Instead of 'a man who sells goods', I may say 'merchant'; instead 'a merchant who sells meat', I may say 'butcher'. In poetry readings, the words of a text cannot be changed; mental processing space can be saved by grouping and clear-cut articulation. Consider again Shakespeare's 'Pity the world, or else this glutton be'. Traditional metrists would say that it begins with an 'inverted first foot'; according to the foregoing argument, a foot cannot be inverted, only the stress can be displaced from a strong position to the preceding weak position. Halle and Keyser would say that such displacement is perfectly acceptable, because being the first syllable in a line, it cannot be preceded by an unstressed syllable.

I handle the situation differently. According to Gestalt psychologists, a closed symmetrical shape takes up relatively little mental processing space. When facing some irregular organization, one must make efforts to organise one of the information streams into a closed symmetrical shape, so as to make room for the processing of the other information stream(s) as well. By displacing the stress from the strong position to the preceding weak position, regularity is upset, and one must

remember that the first syllable is stressed, but occupies a metric weak position; the second syllable is unstressed, but occupies a metric strong position; the third syllable is, again, unstressed, occupying a metric weak position; finally, in the fourth syllable a stressed syllable occupies a metric strong position. This is a lot of information to remember, and may flood the system, so that the rhythmic organisation can be lost. If, instead, the reader recodes this irregular information as a closed symmetrical shape of four syllables (stress valley), s/he may spare mental processing space for the metrical set to reverberate in short-term memory.

Before proceeding, I must briefly clarify another issue, regarding metric boundaries. It is a truism that a line ends with a line boundary. A caesura is a boundary that intrudes into the verse line at the middle. In the iambic tetrameter and hexameter it divides the line into two segments of equal length and equal position structure (each beginning with a weak position and ending with a strong position). In the iambic pentameter it divides the line either into segments of equal length, or of similar structure. If you divide the line into 5+5, you get one segment that begins and ends with a weak position, and one that begins and ends with a strong position. You may obtain segments of similar structures (that begin with a weak position and end with a strong position) if you divide the line into 4+6 or 6+4. Thus, instead of a *point* of caesura, we have an *area* of caesura over positions 4, 5, 6. For reasons related to the limited channel capacity, in a sequence of utterances of unequal length, 'the longest comes last' is felt to be the well-shaped order, as long as no considerations of preference are involved. Consequently, in the iambic pentameter line, caesura after the fourth position is the unmarked one, after the sixth position the marked one. One of the reasons that displacing stress from the second to the first position is more acceptable than in other positions is that in it a stressed syllable in a weak position disturbs regularity, arousing expectations for a return to regularity, followed by a brief period of two positions of uncertainty (when regularity is perceived only as what Seymour Chatman calls the 'metrical set'), and then the two patterns have a coinciding downbeat at a relatively stable point,

before the unmarked caesura, where rhythm becomes fresh and new.¹

As I mentioned, Halle and Keyser account for this differently: displacement of stress from a strong position to the preceding weak position is more acceptable in the first than in other weak positions, because here the stressed syllable cannot be preceded by an unstressed one, and thus no stress maximum in a weak position can occur. However, in positions 3, 5, 7, and 9, too, the stressed syllable in a weak position can be preceded by a stressed syllable in a strong position (as in ‘Of Mán’s first disobedience and the fruit’, for instance), so that no stress maximum occurs in a weak position; still, they are less acceptable than in the first position. Note that if you overarticulate in this line ‘first’ (the deviant stress) rather than ‘Man’s’, you enforce a stress valley that emphatically ends in position six, before the marked caesura, and is perceived as quite acceptable. From this an important principle can be adduced. It’s not enough for an explanation to be logical; it also must be tested against intuitions regarding actual instances.

This model can account for the systematic unequal distribution of stress maxima in weak positions in iambic pentamer lines. The line has three potential boundaries. The stronger the boundary, the stronger its grouping potential. The boundary at the line ending (after the tenth position) is the strongest one; the unmarked caesura after the fourth position is the second strongest; the marked caesura after the sixth position is the weakest. We should expect that stress maxima that start a stress valley that ends in the tenth position should be the most frequent. Here we experience exceptionally great violation of regularity, while regularity is reinstated in the very last strong position, yielding exceptionally great stability and satisfaction. This is, indeed, what we find, since such stress valleys have their stress maxima in the seventh position, where about two thirds of the cases are found. Stress valleys that end in the fourth position, before the unmarked caesura, cannot have a stress maximum at their beginning. Stress valleys that end before the marked caesura in the sixth position entail a stress maximum in the third position; we find there,

¹ Parenthetically I would like to mention another possible reason as well. In ‘Pity the world’, for instance, both the line onset and the stress valley onset are cued by an initial high pitch which, thus, is overdetermined. If the stress valley onset occurs, e.g., in position 3, the initial high pitch is split, and the smooth flow of the intonation contour disrupted.

indeed, nearly one third of the cases. A stress maximum in the fifth position suppresses by definition both the marked and the unmarked caesura in the fourth and sixth position, and ends in the eighth position, which is not followed by any metric boundary. Here we have found, indeed, only very few cases, some of which are rather doubtful. After a stress maximum in the ninth position can come, by definition, no stressed syllable that reinstates metre in a strong position by a coinciding downbeat. Indeed, here I've found only instances that involve emphatic stress, as in 'Bright Star! would I were steadfast as thóu art'.

To test my expectations, I compiled a list of five lines with a stress maximum in the seventh position, such as 'and made him bow to the góds of his wives', 'Burnt after them to the bóttomless pit', and asked members of faculty at the English department of Sussex University to perform these verse lines so as to do as little violence as possible to both the metric pattern and the linguistic stress pattern. Notice that 'to the bottomless pit' would be unmetrical, on different grounds, in the theory of another outstanding generative metrist too, Paul Kiparsky. In his system, a polysyllable that has its stressed syllable in a weak position renders the line unmetrical, even if it involves no stress maximum (according to Halle and Keyser's theory, a stress maximum in a weak position violates metre even if it is a monosyllable). I expected that experienced performers would observe continuity between the unstressed definite article and the ensuing content words 'gods' or 'bottomless'; but, at the same time would suggest discontinuity at the beginning of the required stress valley (how this can be accomplished, see below). I also expected them to overstress the deviant syllable, to foreground the boundaries of the stress valley and to increase the expectation for the reinstatement of regularity, so as to achieve greater satisfaction. The first thing of note was that the performances chose essentially the same solutions, and that those solutions were in harmony with the expectations of my theory. Secondly, my informants were very much surprised to discover that instead of playing down the deviant stress, they rather overstressed it. At that time I had no access to electronic equipment; but later, my instrumental investigations amply confirmed these findings (see below). The fifth item in my list of lines with stress maxima in the seventh position was 'And with these words his temptátion pursued'. While in the preceding lines the stress maxima occurred in

mid-phrase, at the beginning of a word, in this line it occurred in mid-word. I expected that if the informants entered a routine of solution in the first four lines, they would have extra difficulty in this line. And that's exactly what happened.

NG & DNS: One of our interests in intonation comes from its being neither purely sound nor sense, that intonation in spoken performance is interpretive in both the hermeneutic sense literary critics habitually employ, and also the sense of musical interpretation. It seems to us that your work is particularly fascinated in this convergence, and fascinating in its treatment of it. How can our attending to intonation inform our understanding of what poetic 'interpretation' is?

In my 2006 book, "*Kubla Khan*"—*Poetic Structure, Hypnotic Quality and Cognitive Style: A Study in Mental, Vocal, and Critical Performance*, I have discussed related issues at great length, with reference to two excerpts from the last stanza of *Kubla Khan*. Here I can give only a hyperlink to a relentlessly-abbreviated version of that discussion. But before that, to avoid misunderstandings, let me remind ourselves that interpretive statements cannot be true or correct; they can be plausible at best. Where 'a is true' and 'b is true' are incompatible, 'a is plausible' and 'b is plausible' are not incompatible. This is true of our statements about vocal interpretations as well.

[See *Appendix 1* for a development of this point]

NG & DNS: A recurrent feature in your answers here is how you see your thinking as coming out of dialogues with a variety of different theorists of metre, phonetics, and so on...

Apart from Halle and Keyser, two more prosodists had an enormous influence on me, Seymour Chatman and Roger Fowler, both with a strong linguistic bias. Chatman's *Theory of Metre* is a veritable treasure trove of illuminating linguistic analyses and reports on psychological experiments, from which I learnt an enormous lot. However, very frequently he applied those treasures to poetic rhythm in ways that were unacceptable to me. To give just one example, from Chatman I learnt, what may be known to any first-year student of linguistics, that the end of some utterance may be cued by a variety of means: straightforward pause, a falling intonation contour, or the prolongation of the last speech sounds, comparable to *fermata* in music. Usually there is some redundancy, and two or more such cues are

present. It suddenly occurred to me that in enjambment, for instance, where the syntactic intonation contour and the versification intonation contour conflict, they could be indicated by using conflicting cues, at the expense of redundancy. For instance, continuity could be indicated by the absence of a pause, whereas discontinuity could be simultaneously indicated by a falling intonation contour and the prolongation of the last phoneme(s) before the ‘break’. At any rate, that’s how my informants solved the problem of indicating continuity and discontinuity at the same time before stress maxima in a weak position.

Chatman didn’t realise that he had such an effective tool in his treasure trove, by which he could account for the simultaneous presence of conflicting intonation contours. On the contrary, rather, he was a militant exponent of the received view that no more than one intonation contour can be vocalised at the same time. Regarding both metric and intonational patterns, he construes ‘ambiguous’ as ‘capable of alternative actualizations’, rather than *simultaneous* actualizations. In a paper called ‘The Intonational Fallacy’ (1966) he assertively took to task Katherine T. Loesch, who in 1965–1966 proposed non-disambiguating intonation contours for the performance of conflicting intonations (I have noticed that in many issues involving ambiguity Chatman determinedly preferred the non-ambiguous option).

NG & DNS: Yes, we recall that in your essay for *Thinking Verse II*, you argued that it is possible for a reading to register two conflicting intonation contours at once. This will seem counterintuitive to many of us, or at least anomalous, if simply because we have only one voice. How did you come to this conclusion, and how do you account for it?

I argue there, and in many other places, that the conflicting cues distort to some extent the intended phonological entity, and that our ability to handle such conflicting cues is based on our phonetic competence, an ability characterised by Philip Lieberman in this way: ‘the listeners perceive these signals by means of a feedback mechanism of the analysis-by-synthesis type, in which they use their knowledge of the phonologic features that produce intonational signals.’² We may assume that in the rhythmical performance of poetry this mechanism is turned to

² Lieberman, *Intonation, Perception and Language* (Cambridge, Mass.: MIT Press, 1967), p. 60.

an aesthetic end: listeners decode the deviant articulatory and intonational signals in terms of the phonological and versification features whose interaction produced them. Iván Fónagy speaks of double-encodedness. I assume triple-encodedness: the same noises may convey some phonological entity (phoneme or intonation contour, etc.), some emotive attitude and some versification structure (in psychology, Wilma Bucci speaks of multiple encodedness).

Loesch's and my view appear to be a minority view. In my book *Poetic Rhythm* I quote what appears to be the 'canonic' view from the 'Performance' entry of *The New Princeton Encyclopedia of Poetry and Poetics* (1993) which, in turn, quotes Seymour Chatman on ambiguity. 'Chatman isolates a central difference between the reading and scansion of poems on the one hand and their performance on the other: in the former two activities, ambiguities of interpretation can be preserved and do not have to be settled one way or the other ('disambiguated'). But in performance, all ambiguities have to be resolved before or during delivery. Since the nature of performance is linear and temporal, sentences can only be read aloud once and must be given a specific intonational pattern. Hence in performance, the performer is forced to choose between alternative intonational patterns and their associated meanings'.

NG & DNS: And how were you able to set up a position opposed to Chatman?

I received instrumental support for my position from an unexpected direction. In the late nineteen eighties I met Tom Barney at an empirical literature studies conference in Budapest. He was at the time an MA student at Lancaster University. Tom's lecture, extracted from his master's thesis, did exactly what Fry told me was impossible to do. By the same token, without ever having heard of my work, he proved that my insight regarding conflicting cues was workable on the computer. I decided that whoever taught him would be good enough for me too. For my next sabbatical I went to Lancaster University, where I had the benefit of Gerry Knowles's guidance and Tom Barney's assistance.

As a result of this work, my list of cues for discontinuity was extended with three additional major items, an overarticulated **stop release** (an audible burst indicating

the end of the occlusion that articulates a stop consonant) at the end of the line before the enjambment where applicable, or a **glottal stop** before the line following the enjambment, where applicable. A word cannot begin with a vowel, it begins with a glottal stop, but in connected speech it usually disappears. When you say 'I didn't say 'a name', I said 'an aim', you disambiguate the phrase by reinserting the glottal stop before 'aim', without inserting a pause. The third one is **late peaking**; this concerns the alignment of the intonation contour with vowels and consonants. Knowles writes:

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: p. 294)

Knowles calls this point the *accent point*. 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. 'Peak position would seem to be a continuous variable' (*ibid*). Early peaking presses backward, away from the break; late peaking presses forward. If it occurs after the caesura, it presses away from it; if it occurs before the caesura, it presses across it. Even as a teenager I discerned some mysterious forward push in some readings at certain points. All the repertoire of cues for continuation and discontinuation couldn't account for that mysterious push. When I began my empirical explorations in the performance of poetry, I started with a widely-entertained notion that phonemic categories can be conveyed by a variety of phonetic cues, but what matters is the phonemic category, not the specific cue (voicing of a consonant, for instance, can be cued by straightforward activation of the vocal chords, or by lengthening the preceding vowel; what is of importance, however, is not the specific cue, but the fact of voicing). When I read Gerry's paper on late peaking, everything suddenly became clear. And the confirmation was visible on the computer screen. I had to realise that many of the rhythmic effects are determined not by the phonemic categories but the pre-categorical phonetic and acoustic cues.

For our present interest it is important that peak position may affect the grouping of syllables. Knowles suggests the possibility that behind the phonological contrast of tone there is a functional contrast between an ‘initial’ marker and a ‘final’ marker. ‘Looking at all final tones before different kinds of boundary it was found that the greater the boundary the earlier the peak, with the earliest peaks preceding the silent pauses forming a major tone group boundary. The position with regard to initial position is not so clear: the latest peaks follow [...] pauses [...] within a major tone group [...]. Tones at the beginning of a major tone group are neither early nor late’ (Knowles, 1992: p. 296). I have found that late peaking is much more widespread in poetry reading than in ordinary speech, and not only positioned as specified by phoneticians. I have also found that the poetic effect of late peaking can be accounted for by a finding of Gestalt psychologists. If an event intrudes in a perceptual whole at the middle, it enhances its sense of stability. If it intrudes between the middle and a boundary, it is perceived as pushing toward the boundary. Gestaltists call this *perceptual forces*. The further away the intruding event from the middle, the stronger the perceived force. I have found that this is an exceptionally effective means for the reciter to manipulate cues to generate a rhythmical performance as defined above.

The great breakthrough occurred when I realised that the same conflicting vocal cues as found in enjambment could be used for saving mental processing space for the rhythmical performance of additional deviant configurations: consecutive stresses, and stress maxima in weak positions as well. (Luckily enough, this happened before Christmas, because when I came home for Christmas vacation, I had an emergency operation and could not return to Lancaster anymore. But then I could already pursue my research on my own.)

NG & DNS: Could you give an example?

The following verse line by John Donne, with its three readings illustrates two crucial issues: how a problematic stress valley beginning with a stress maximum in the seventh position is performed by experienced readers; and how an intonation contour ending with a minute boundary intonation known as ‘fall-rise’ can be decoded by listeners in terms of the conflicting phonemic pattern and versification

pattern whose interaction generated it.

7

Buffet and scoffe, scorge, and crucifie meel!
Donne, 'Holy Sonnet XI' l.2

This line would be 'unmetrical' under at least two different criteria, propounded by the various kinds of generative metrics. According to Halle and Keyser, a stress maximum in a weak position renders a verse line unmetrical, even if no polysyllable is involved; according to Kiparsky, the stressed syllable of a polysyllable in a weak position renders the line unmetrical, even if it is not a stress maximum. Now *cru-* in the seventh position is both. Nonetheless, experienced performers seem to be in agreement as for the performance pattern that may render such a line rhythmical: a stress valley comprising the last four syllables. The performers whose performances are to be discussed ('TB' and 'JH') were at the time PhD candidates at the Department of Linguistics and Modern Languages at Lancaster University.

It will be noticed that all three performances discussed below conspicuously overstress rather than downplay the deviant stress. This results from the interaction of the cue for phonemic stress on *crucifie* and the need to foreground the onset of the performance pattern 'stress valley', required by the 'violation' of metre. Intuitively, in many performances it is an exceptionally high pitch that sets the syllable *cru-* apart from *and*, pushed forward, as it were, by the perceived leap to the high pitch, occasionally corroborated by a late peak. Such a late peak generates discontinuity with the preceding syllable; by the same token it groups the syllable with the ensuing syllables that constitute the stress valley. Indeed, in TB's reading the intonation contour on *and* rises from 120.492 to 121.823 Hz, then on *-cru* it resets to 156.383 Hz; then it falls to 117.287 Hz on *-ci-*. Owing to the prominent pitch discontinuity, the long-falling contour does not indicate here closure, but the start of a new movement in mid-phrase.

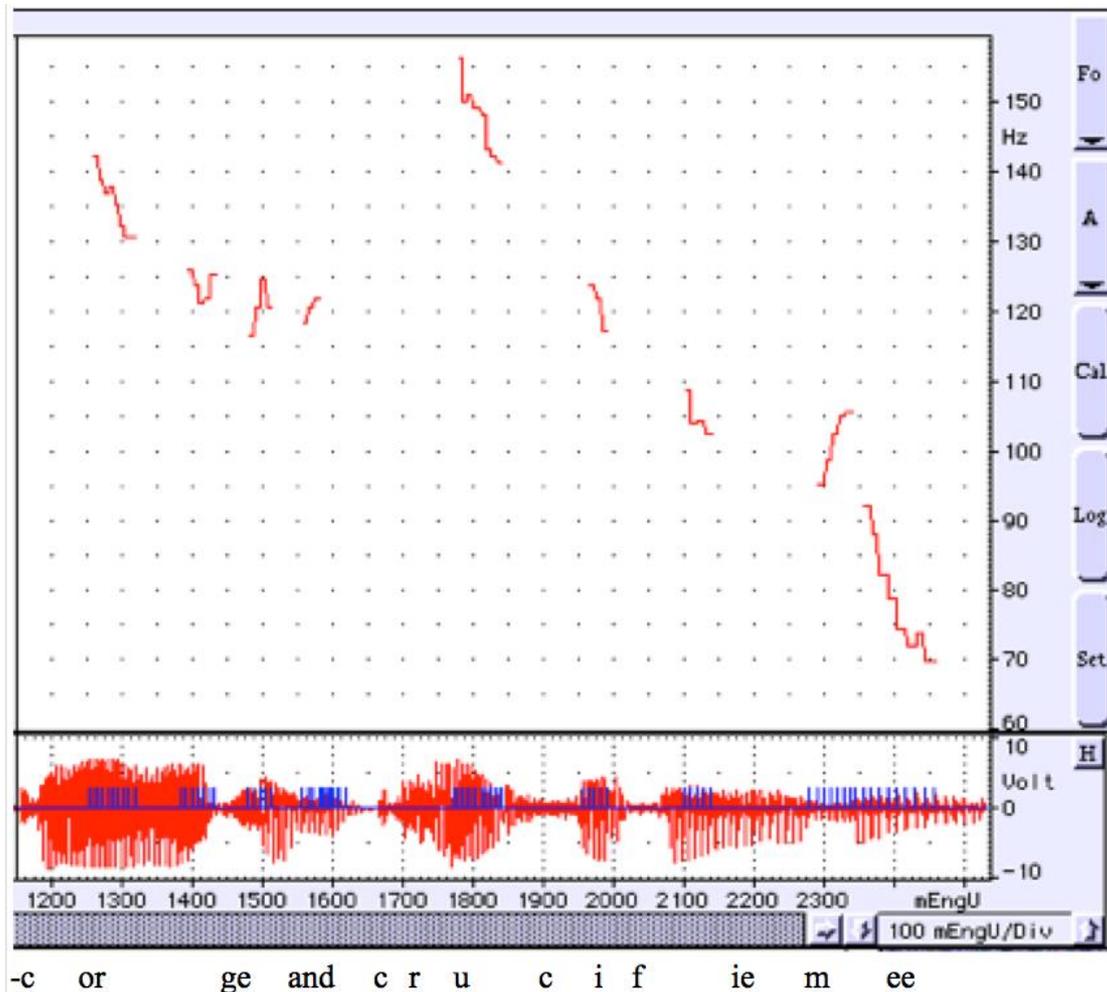


Figure 1 Wave plot and F₀ extract of '-corge and crucifie mee' in TB's reading. **Listen**

In another delivery instance (by JH), exactly the same sort of stress valley is discerned ('crucifie mee'); again, it is segregated from the preceding conjunction by an obtruding instance of pitch discontinuity (the jump to /u/, from which the intonation pattern descends gradually). The contour of *and* falls from 87.849 to 79.603 Hz; on *-u-* pitch resets to 117.287 Hz; then, after some curling, returns to virtually the same peak (117.129 Hz), persisting until slightly after the middle of the vowel (this is what Knowles 1992 calls late peak, and according to Gestalt theory may be expected to exert a 'forward-pushing *perceptual force*'; see above). This seems to be just enough to suggest pitch discontinuity and some forward grouping of *cru-*

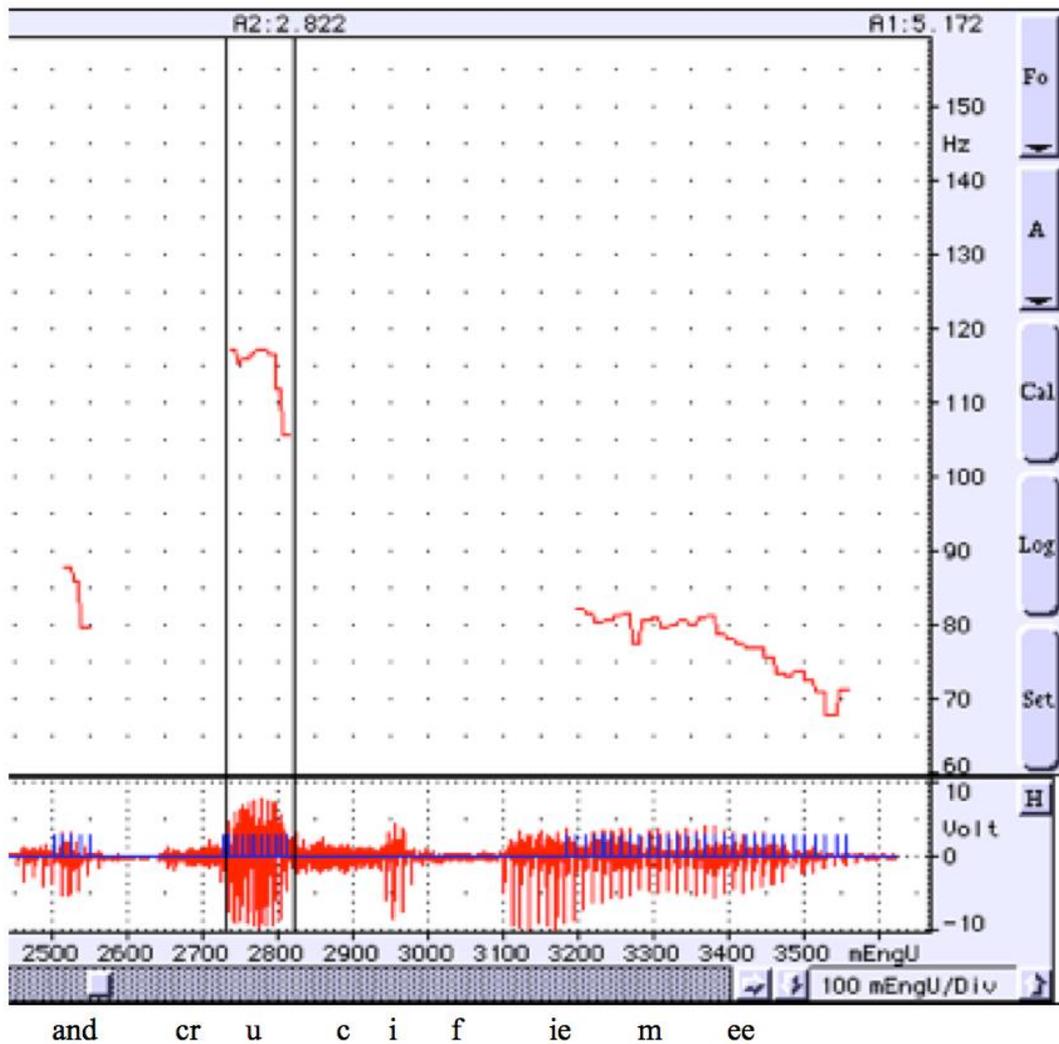


Figure 2 Wave plot and F₀ extract of JH's first recording of 'and crucifie mee'. Listen.

JH recorded the verse lines twice: once in context, and once in isolation. In the latter recording (Figure 3), with reference to *cru-*, the peak is quite early in the vowel; but he creates a clearly-noticeable pitch discontinuity, by a big upward leap between the two words (the pitch of *and* swerves between 90.000–84.160 Hz, ending at 88.200 Hz, from which it leaps to 123.184 Hz, where a downward movement begins, falling to 92.259 Hz on *-ci-*).

In Figures 1 and 3 one may notice another important point. Not only is *and* discontinuous with *cru-*, but it is also conspicuously continuous with *scorge*: the

release of the affricate is run into the vowel of *and*; and the intonation of *and* is in the pitch region to which the long intonation contour of *scorge* falls (still, they are perceptually segregated by the long terminal contour articulating the boundary of *scorge* and the exceptionally strong release of the affricate; indeed, the stressed syllable conflicting with the weak position it occupies requires exceptionally strong articulation of the word boundary—to save mental processing space for the perception of the conflicting patterns). Moreover, in Figure 3 the powerful release of the affricate in *scorge* is preceded by a longish /r/. This further enhances its forward push toward *and* but enhances the release too, indicating discontinuity.

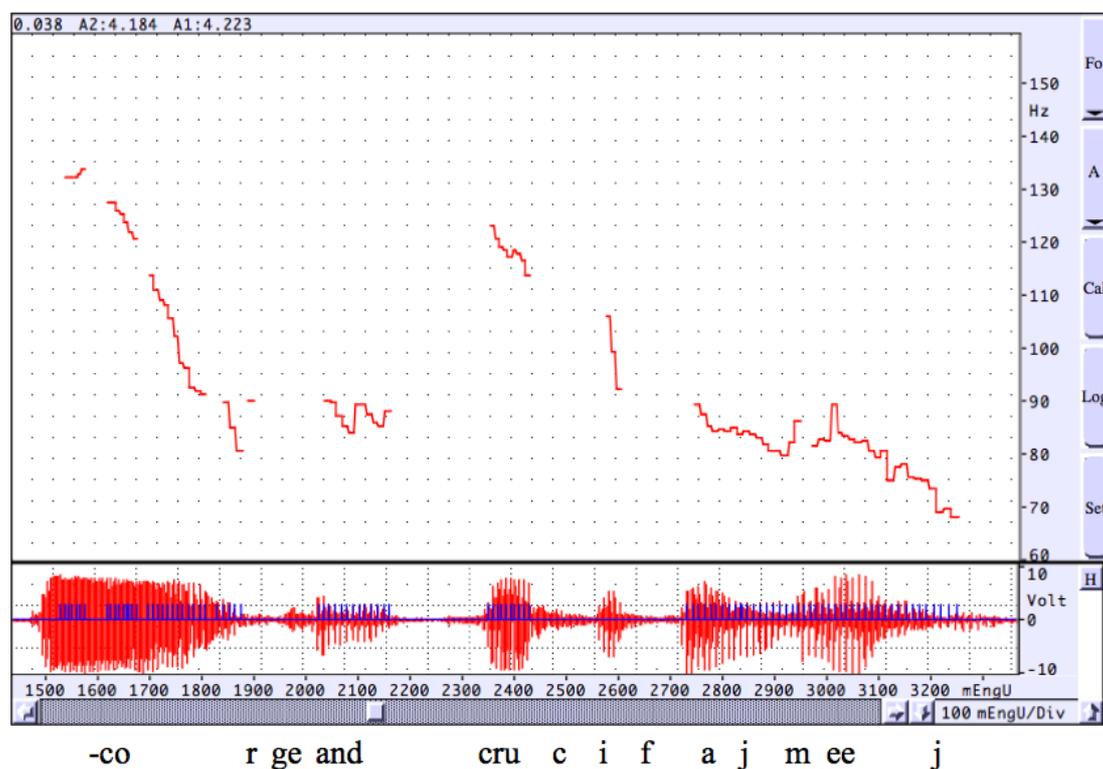


Figure 3 Wave plot and F₀ extract of JH's second recording of 'corge and crucifajm ee j'. **Listen.**

The reciter has another problem here: the last syllable of this line bears no content-word stress, whereas the last position of the line requires confirmation by stress; what is more, the stress valley demands that this stress be prominent indeed. Both

TB and JH made a remarkable effort to face this challenge by a compromise between stressing and not stressing, but in different ways. In a reading in which *mee* is unstressed, one might expect the intonation curve to fall on, or immediately after, *cru-*, and keep the rest of the pitch sequence level at the base line. In a reading in which *mee* is stressed, one might expect the intonation curve to rise on *mee*. In TB's reading neither of the two happens: the tone falls throughout *crucif*e**, to 102.558 Hz; then, on *mee* it slightly moves up from 98.879 to 105.502 Hz, and then falls down to 69.778 Hz, and is perceived as an 'allophonic' stress. In JH's first reading, pitch falls moderately on *-fie mee* and then, contrary to expectation, it rises again a few Hzs, yielding a minute fall-rise boundary tone. Such a small rise usually goes unnoticed; here it is clearly perceptible, probably because a fall is expected. In his second reading there is a downward jump of intonation from *cru-* to *-cif*e**. Normal expectation would be for a downward continuation on *mee*. Contrary to expectation, pitch falls on *-fie* from 89.271 to 79.603 Hz, and then rises again to 86.133 Hz; and then rises on *mee* from 81.667 to 89.271 Hz, from where it has a long fall to 68.056 Hz.

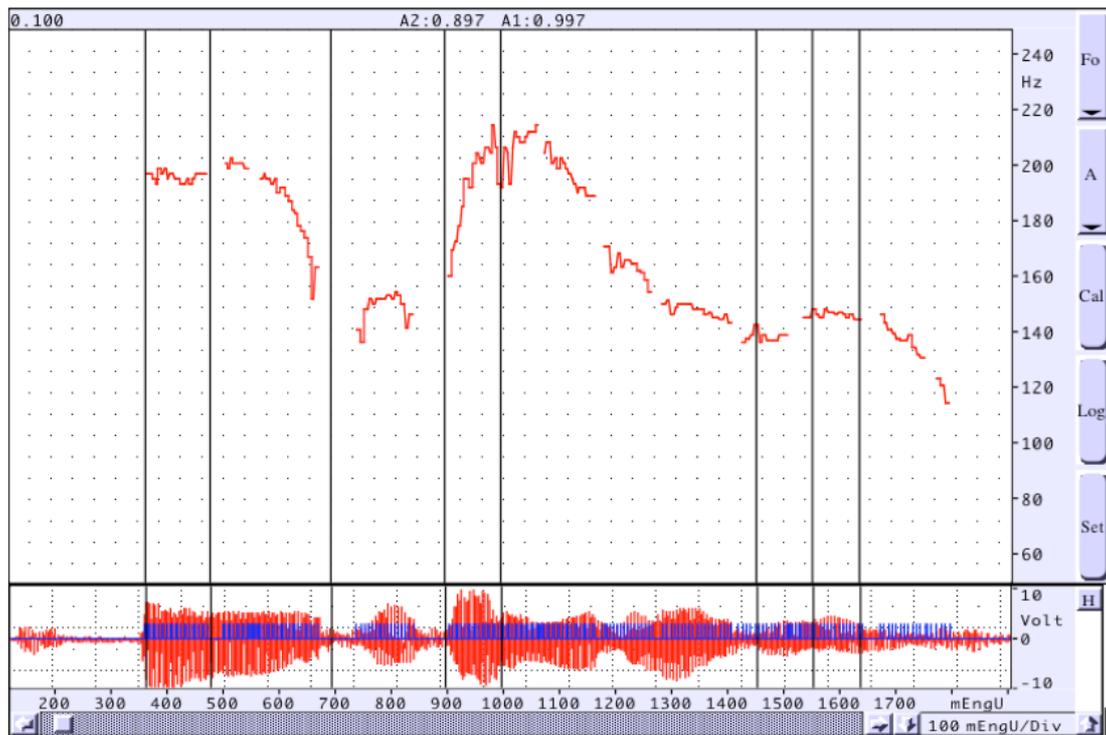
In the first reading, this is reinforced by an uncommon manipulation of articulation. When isolated, the first vowel of the diphthong in *crucif*e** is perceived as reduced; the vowel of *mee*, by contrast, is turned into a diphthong, with an audible [j] at the end. In the flow of reading this is not perceived. Together with the minute rise of pitch, it is merely sufficient to indicate that *mee* is more prominent than *-fie*, without being perceived as a stressed syllable. In the second reading, only the vowel of *mee* is turned into a diphthong.

The two performers accomplished, then, in the last position a mission impossible: to suggest a complex intonation contour that could be decoded as indicating, at the same time, as metric downbeat and the absence of linguistic stress in the tenth position.

NG & DNS: How can we account for such a phenomenon?

Let us turn to the Marlowe Society's reading of Shakespeare's Sonnet 2, line 7:.

To say within thine own deep-sunken eyes



to s a y w i t h i n t h i n e o w n
 Figure 4 Wave plot and F⁰ extract of ‘to say within thine own’ in Marlowe Society’s reading. **Listen**

The markers indicate diphthong and vowel boundaries. Notice the late peaks on *say*, *-thin* and *own*.

Here we encounter something quite unusual. The second syllable of ‘within’ bears no lexical stress at all. Yet it displays an exceptionally high intonation leap. Yet, Simon Callow too observes a leap on this syllable. This intonation contour is required here as one possible solution to a versification problem: the assignment of caesura is quite problematic.

One may note, in addition to the high leap a change of direction in the intonation contour, an exceptionally long /n/ (its length is roughly equal to that of ‘thine’) and late peaking. The combination of these vocal cues results in more than usually clear-cut articulation of the last syllable before the caesura. Apparently, metric regularity in the first six syllables ‘To say within thine own’ is straightforward enough, and there would appear to be no perceptual problem to justify such extravagant devices.

But the reciter has a real problem: there is here no natural point for a caesura. Positions 4 and 5 are occupied by function words ('within *thine*'), and the content word in position 6 (*own*) must be grouped forward emphatically, for a variety of syntactic and metrical reasons. Thus the reciter has the challenge of suggesting continuity and discontinuity at the same time. Continuity is demanded by syntax; discontinuity by the need to observe a caesura. Discontinuity is indicated here by the falling intonation contour, and the unusual lengthening of the /n/. As in *fermata* in music, such lengthening suggests the perceptual end of some unit—the lack of onward movement.

Shakespeare's line is very complex in many ways. I shall focus only on two points in it: there should be a caesura, that is, discontinuity, after a function word in the middle of a metrically and syntactically complex phrase, and there should be discontinuity after the reporting phrase 'to say', that might sabotage the sense of discontinuity at the caesura in mid-phrase. The reciter effects continuity and discontinuity at the same time, in both cases. There is no measurable pause at either point. This takes care of continuity. The exceptionally long /j/ (in 'say') and /n/ (in 'within'), coupled with a falling contour on both periodic continuants, take care of discontinuity. Now by careful listening one may discern a forward push (across the non-existent break) at both points, caused by late peaking. In 'say' there is a peak on the second sound of the diphthong; in '-thin' there are two equally high late peaks at the end of /i/ and after the vowel, on /n/.

What I have aimed to do in these readings and elsewhere is demonstrate how the same noises may cue at the same time phonemic, emotive and versification patterns. The perceived shape is generated by the interaction of two or more patterns. I suggested that the listener decodes those shapes in terms of the patterns whose interplay generated them and responds to them differentially. Such ability of handling conflicting or deviant cues seems to be built into our phonetic competence and is turned to an aesthetic end.

NG & DNS: It strikes us that intonation, for you, attains particular importance here—namely, where phonetic competence and the aesthetics of performance converge...

In principle this is true, but here aesthetics is limited to versification pattern. Let

me illustrate my conception with one of my standard examples. Consider the following verse instance from Keats's 'Ode on a Grecian Urn' in which the versification unit (the verse line) conflicts with the syntactic unit (the clause), that is, when the phrase or clause runs on from one line to the next one. In my book *Poetic Rhythm* I compared two recordings by two leading British actors, Douglas Hodge and Michael Sheen.

Sylvan historian, who canst thus express
A flowery tale more sweetly than our rhyme...

The overwhelming majority of listeners made the judgment that Hodge offers a rhythmical solution to the problem, by suggesting continuation and discontinuation at the same time at the end of the word 'express'. In Sheen's reading, by contrast, 'A' at the beginning of the next line is irritatingly continuous with 'express'. There is no measurable pause in either of the readings between the two words; and this takes care of syntactic continuity. Two significant differences between the two readings may account for the perceived difference between them. First, in Sheen's reading the /s/ of 'express' is inseparably run into 'A', whereas in Hodge's reading we may discern a glottal stop that perceptually separates the two words, indicated by a minute 'lump' in the wave plot. Second, the syllable '-press' in general, and the closing /s/ in particular, are considerably longer in Hodge's reading than in Sheen's. (see Figures 5–6).

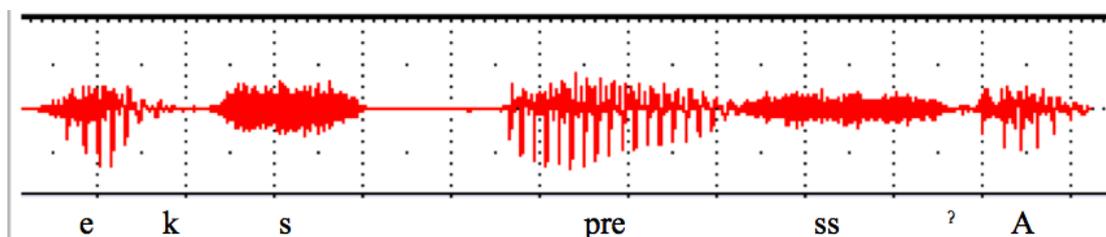


Figure 5 Wave plot of 'express / A' in Hodge's performance (? indicates glottal stop).

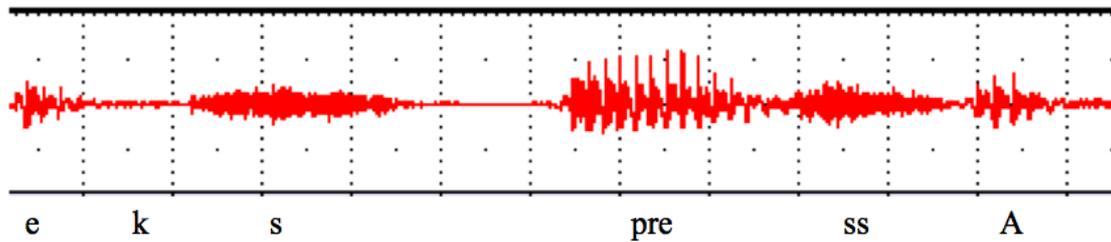


Figure 6 Wave plot of 'express A' in Sheen's performance (no glottal stop).

Further experimentation suggested that by electronically manipulating Sheen's reading one may reproduce in it the same perceived effect as in Hodge's. I copied a section of Sheen's /s/ and repeatedly re-pasted it, prolonging the /s/. Then I copied from Hodge's reading the glottal stop and pasted it into Sheen's reading, before the A'. Again, the majority of listeners judged that in the doctored, but not the original, version conflicting cues for continuation and discontinuation are provided.

In addition, Sheen overstresses 'our', violating the iambic cadence of the line, by generating a virtual stress maximum in the ninth position (the fact that 'our' is an unstressed function word whereas 'rhyme' is a stressed content word becomes practically insignificant). The easiest way to handle such a case is to say that 'our' should not bear emphatic stress, and there's an end to it. My way of handling it is to ask whether a performance can be imagined or secured in 'our' bears emphatic stress, while the iambic cadence is still preserved.

NG & DNS: So it is an evaluative question, rather than (a) purely descriptive or (b) definitional?

In principle, it is strictly descriptive. But if you happen to believe that in spite of any deviations the verse line should perceptually be integrated in the iambic context, it becomes evaluative. I personally believe so. I believe that in spite of far-reaching permissiveness, somewhere some constraints must be observed. Otherwise, anything goes, and art stops being art. Nobody can compel you to perform a verse line rhythmically. But, equally, one must find out whether a performance *can* be imagined or secured in which 'our' bears emphatic stress, while the iambic cadence is still preserved—just in case someone wants to perform it rhythmically. Doctoring

the recording of a verse line could serve as evidence for a positive answer.

I made an attempt electronically to restore the iambic cadence without detracting from the emphatic effect. In the second manipulated version I used the application Audacity to reduce the tempo of the word ‘rhyme’, lengthening it without affecting pitch, so as to make it break rhythmically even with the emphatic stress on ‘our’. This doctoring exercise provides evidence against the equal (or proportional) timing theory: extended duration in the last position broke even with stress indicated by pitch, not duration, in the last-but-one position. Perceptually, the increase of stress in the last but one position sabotages the closure of the line by a coinciding downbeat in the last position, and the line remains open-ended. According to Christine Bartels, ‘statements bearing ‘declarative *fall*’ [...] convey a sense of self-containedness, closedness, “finality”.’³ It is the effect she attributes to falling intonation contours that seems to compensate for such open-endedness: ‘statements bearing ‘declarative *fall*’ convey a sense of self-containedness, closedness, ‘finality’. Prolonging that *fall* seems to enhance this effect. Here the ‘closedness, finality’ effect of intonation is turned to an aesthetic end. From the syntactic point of view, it typically enhances the conclusiveness of the state of affairs suggested by the meaning; here it is transferred to the verse line and serves to indicate its completeness. By the way, this is exactly how experienced performers render stress maxima in the ninth position rhythmically acceptable.

NG & DNS: What is fascinating here is how technological advances in speech analysis allow us to answer, or at least reconceptualise, long-standing issues. How do you see the future for such technologies in our analyses not just of individual performances, but of a text’s possibilities, and of how we theorise metre more generally?

While I perfectly agree with your comment, my record in foretelling the future is rather poor. Speaking for the present, I have emphasised throughout the discussion that I don’t presume to prescribe how a verse line should be performed. Still, I insist time and again that one must find out whether at least one performance *can* be imagined or secured that conveys the text’s complexities—just in case someone

³ Christine Bartels, *The Intonation of English Statements and Questions: A Compositional Interpretation* (New York & London: Garland Publishing, 1999), p. 59.

wants to render them in a performance. Doctoring the recording of a verse line could serve as a most convincing evidence for a positive answer. I must, however, add a caveat. Such doctoring exercises are fascinating; but in many instances the output of the computer merely re-states in the visual medium and confirms what you *hear* anyway.

NG & DNS: So the technological advances only reinforce the importance of the human ear for discerning rhythmic pattern and variation?

I would prefer ‘confirm’, or ‘verify’ to ‘reinforce’. The doctored version enables the human ear to directly experience what I claim theoretically. It is an empirical proof of the psychological reality of the solution.

NG & DNS: You mentioned Fowler as well as Chatman earlier...

Yes. Among other things, it was Fowler's paper ‘Prose Rhythm and Metre’ that directed me to Wellek and Warren's threefold distinction between Prose Rhythm, Metre, and Performance. Of his many brilliant insights I will mention only one.

Centuries of prosodists in many languages have quoted with approval Aristotle's statement: ‘The iambic is the characteristic rhythm of people as they talk’. The issue here is Why is the iambic perceived as more natural than the trochaic in a wide variety of languages? At the end of his “‘Prose Rhythm” and Metre’, Roger Fowler comes to the opposite conclusion:

This, paradoxically, may help to explain why the iambic measure is felt to be suited to English: not because its pattern corresponds to the prose rhythms of language, for it does not; but because it necessitates a constant syncopation of prose rhythm against its own rhythm, inviting poets to be metrically complex, not to jog along with simple regularity.⁴

Surprising as this may be, the great variety of languages to which the iambic measure seems to be ‘suited’ (even to Hungarian!—where stress invariably falls on the first syllable) supports this conclusion. I would, however, prefer ‘tolerates’ to ‘necessitates’. Verbal *necessity* alone cannot account for syncopation; it cannot

⁴ Roger Fowler, “‘Prose Rhythm” and Metre’, in Fowler (ed.), *Essays on Style and Language* (London: Routledge & Kegan Paul, 1966), p. 99.

explain why Pope should resort to syncopation less frequently than Milton (as Fowler himself asserts), or why Shakespeare in his earlier work should resort to it less frequently than in his later work. It seems, rather, that syncopation is related to wider aesthetic issues (what I have called above convergent and divergent styles) and that it takes deliberate daring to abandon the security of established strong shapes. Far from being a necessity, syncopation is a daring achievement; it is a delicate balance between prose rhythm and metre.

Chatman too acknowledges this difference between the iambic and the trochaic. ‘The trochaic mode more easily violates normal prose accentual patterns; it quite insists on dominating the rhythm. Iambic verse seems not to exert its will so rigidly’ (p. 141). As an explanation for this peculiar character, Chatman quotes Martin Halpern, and then offers to substitute his own explanation:

Halpern’s thesis is that trochaic verse, along with anapestic and dactylic, is a subspecies of the native Germanic strong-stress verse, which he feels is both isoaccentual and isochronic [...] I think his observations about the relative inflexibility of the trochaic verse are correct, but would suggest another cause, namely the comparatively short history of the mode. The sophisticated smoothness of the iambic verse has been long in developing; trochaic verse, however, was not taken very seriously in England until the nineteenth century. (p. 141, n.)

Both explanations have the same weakness: neither the Germanic strong-stress verse, nor the history of the English metre, are *perceptual features* of an actual piece of trochaic poetry. As Coleridge put it, ‘nothing can permanently please which does not contain in itself the reason why it is so, and not otherwise’ (*Biographia Literaria*, Chapter 14). A child who attends to ‘Tackle, tackle, Mother Goose,/ Have you any feather loose?’ responds to the peculiar trochaic quality in the poem with no particular instruction in the history of English and Germanic verse.

I, again, attempted to handle the issue differently, and in a way that contains ‘in itself the reason why it is so, and not otherwise’. I relied on two sets of experiments, one by D. B. Fry, the other by H. Woodrow. Curiously enough, again, both experiments can be found in Chatman’s book (he even replicated Fry’s results by other means), but, again, they didn’t affect his argument in this respect.

One would imagine that in an endless series of equidistant *tick-tacks* no preference would be given to iambic or trochaic rhythms; the only distinction should be whether the series began with an upbeat or a downbeat. Experimental psychology, however, shows that this is not so. H. Woodrow found in his tick-tack experiments, back in 1911, that in a series of tick-tacks, ‘with equal temporal spacing, a regularly recurring, relatively greater intensity exerts a group-beginning effect, and a regularly recurring, relatively greater duration a group-ending effect’ (Woodrow, 1951: p. 1233). ‘Intensity has a group-beginning effect: duration, a group-ending effect: pitch, neither a group-ending nor a group-beginning effect’ (Woodrow, 1911: p. 77). Eighty years later, in Chapter 5 of his doctoral dissertation, Curt Rice (1992) replicated Woodrow’s experiments: ‘The technological resources for conducting this research are dramatically more sophisticated than those which Woodrow had available’. At variance with Woodrow, Rice (1992: p. 198) showed that variations in pitch do lead to a significant shift towards iambic groupings.

Now, coming back to verbal rhythms, spoken language consists of syllables of varying duration. Schramm (1935), Fry (1958) and Chatman (1965) have demonstrated that the acoustic cue to linguistic stress is a complex of pitch, duration, loudness, in this order of decreasing effectiveness. Usually, at least two of these three components are present. If pitch differences are irrelevant to grouping direction, and duration differences are more effective in stress perception than amplitude differences, end-accented metres should be more natural in poetry in a variety of languages. If variations in pitch also lead to a significant shift towards iambic groupings, it should reinforce this effect. Significantly, even in Hungarian poetry, where stress is invariably on the first syllable, the iambic metre is far more natural and widespread than the trochaic.⁵

⁵ A short time after my encounter with Jay Keyser, Eddie Epstein, editor of *Language & Style* and co-author of the Epstein and Hawkes metric theory invited me in London for a meeting, following the rejection of a paper of mine. We discussed my theory for several hours. He was especially interested in my discussion of the differences between the iambic and trochaic metre, and the special status of the iambic. He asked me to send him the paper I was writing; but, he added, it will have to undergo the usual refereeing process. I was very much disappointed when he informed me that he had been advised to reject my paper. I was even more unpleasantly surprised when, a few years later, *Language & Style* published a paper addressing precisely those

NG & DNS: You have mentioned your dialogue with the work of Halle and Keyser, Seymour Chatman, and Roger Fowler. What is the relationship of your work to Marina Tarlinskaja's? You both explore the effects of sometimes the same metric configurations, even discover similar qualities, but your explanations seem to rely on very different assumptions. And your general approaches, cognitive and historic, respectively, seem to be opposite to one another.

I have actually answered these two questions at great length, one published and one unpublished. In spite of our great theoretical differences, Marina and I are old and very good friends. On one occasion, the editor of *Poetics Today* gave me the report of an anonymous reviewer on my 1996 paper 'Rhyme and Cognitive Poetics'. The reviewer warmly recommended publication, and then added six pages of criticism of cognitive poetics. After reading the first paragraph, I said that there is only one person in the world who could write this: Marina Tarlinskaja. He refused to assert or deny it. I answered the criticism in a longish postscript to that paper which, eventually, we decided not to publish. Later, Marina admitted having written the report, and we decided to publish together a signed controversy. However, alas, after several attempts it didn't work out and we gave up.

[See *Appendix 2* and *Postscript*]

NG & DNS: Your work theorises rhythm through analyses of recordings of different kinds of readers (poets, actors, students, etc.), and so it would seem far removed from the tradition that treats rhythm as either an ideal template, on the one hand, or as purely sensuous on the other. This leads us to wonder about whether rhythmic analysis should be descriptive (simply analysing individual readings), or normative (how one determines what makes a 'good' reading), or productive (accounting for possible, and future, readings, as well as actual ones). How do you see these forms of analysis as relating in your own work?

questions, relying on precisely Fry's and Woodrow's experiments. I didn't ask the questions that suggested themselves; only in the Preface to my *A Perception-Oriented Theory of Metre* (1977) I mentioned Robert P. Newton's essay 'Trochaic and Iambic' (*Language & Style*, Vol. 8, No.2, Spring 1975, pp. 127-156), 'which makes a few points uncannily similar to mine in the chapter 'Some Remarks on the Nature of Trochees and Iambis''. (By the way, I am now reprinting updated version of the paper in my book forthcoming with Oxford UP, *Poetic Conventions as Cognitive Fossils*.)

I have never understood the tradition that treats rhythm as ideal, or non-perceptual. I have no logical arguments to prove that poetic rhythm is a perceptual phenomenon. But when I listen to poetry readings, or when I sub-vocally perform a poem, I do *hear* the rhythm. And I happen to know a few more people who claim to hear poetic rhythm. There seems to be some confusion between metre and rhythm. Iambic pentameter means that there is a metric foot consisting of an unstressed and a stressed syllable; and there are five such feet in a verse line. In the first 165 lines of *Paradise Lost* there are exactly two such lines. Yet, as I said, Milton is *heard* to be one of the most musical poets in English. A similar story can be told about Shelley. In my conception, metre is an abstract pattern, but rhythm is *heard*. The question is what is the mode of existence of this rhythm.

Metre has, in spite of all, considerable psychological reality. To return to lines already mentioned, lines such as ‘Tackle, tackle, Mother Goose’, or ‘Tyger, Tyger, burning bright’, the metric and the linguistic stress pattern coincide, so that both are immediately observable. In other verse lines, where metre is suspended in the acoustic signal for a brief period of time, it may exist as an expectation, or as an acoustic trace (what Chatman calls ‘metrical set’), or both, as in

Pítý the wórld, or élse this glútton be
w s w s w s w s w s |

Here a stressed syllable intrudes in the first (weak) position upon metric regularity, arousing expectations for a return to regularity; then in the next two positions metre is suspended in the acoustic signal, then reinstated in the fourth position where the two patterns have a coinciding downbeat, making rhythm fresh and new, precisely before the caesura. When regularity is suspended in the linguistic dimension, the regular alternation in the versification pattern may reverberate for a very brief time in short-term memory. The stress pattern of the first four syllables in this line constitutes a closed symmetrical shape (two stressed syllables enclose two unstressed ones; alternatively, the second half is the mirror image of the first half), saving mental processing space for the simultaneous perception of the acoustic signal of the linguistic stress pattern and the suppressed metric pattern reverberating

in acoustic memory.

I have adopted from Wellek and Warren the assumption that in order to account for poetic rhythm, one must assume three dimensions: versification pattern (metre), linguistic pattern ('prose rhythm') and pattern of performance. Wellek and Warren emphasise that when one makes measurements in a recording, one is not making measurements in the poem but merely in a casual performance. To reiterate: I don't conceive of performance as of mere vocalization of a written poem; I conceive of it as of a problem-solving activity: a rhythmical performance accommodates the conflicting linguistic stress pattern and the versification pattern so that both are perceived simultaneously. Thus, a rhythmical performance is not a casual vocalization but is constrained, not determined, by the conflicting patterns of language and versification, as well as by the relevant cognitive apparatus. In this way it may be productive, that is, accounting for possible and future readings, as well as actual ones.

Such problem-solving activity is comparable to the understanding of metaphors, where conflicting meanings are accommodated in a semantic interpretation. Thus the performance of poetry is governed by a homogeneous set of rules. Generative metrists have reinvented Wellek and Warren's first two patterns, but ignore the pattern of performance (some explicitly reject it).

You mentioned the tradition that treats rhythm as ideal, or non-perceptual. Well, some modern prosodists even came to the conclusion that English blank verse is often verse only for the eye, invoking Dr. Johnson (e.g., Fraser 1970). But Dr. Johnson knew better than that: '... there are only a few skilful and happy readers of Milton who enable their audience to perceive where the lines end or begin. 'Blank verse,' said an ingenious critick, "seems to be verse only to the eye.'" (*Lives of the Poets*). Significantly, Fraser quotes not all of the excerpt, only the quotation from the 'ingenious critick'.

Indeed, the surrounding empty space that indicates line endings in printed verse is not available in vocal performance. I claim, however, that just as white spaces break up the series of black marks on the paper into smaller perceptual units whose end may or may not coincide with the end of syntactic units, in aural perception, certain vocal devices may break up the text into versification units, and even indicate

conflicts of versification and syntactic units. As Dr. Johnson said, there *are* a few skillful and happy readers of Milton who enable their audience to perceive where the lines end or begin. Just as the graphic arrangement on the page presents the lines as perceptual units to the eye, the intonation contours heard in the reading of poetry present the lines as perceptual units to the ear. In fact, the main function of the graphic arrangement on the page is to give the reader instructions concerning the intonation contours appropriate to the lines; whereas the intonation contours appropriate to the syntactic units are determined by our linguistic competence.

NG & DNS: So would you agree with the definition of enjambment as the disjunction not of verse line and syntax, but verse line and intonation contour? Or do you see intonation contours as effectively mapped onto syntax?

I think the latter is the case.

Obviously, enjambment generates some ambiguity, generally deemed as a good-making feature of poems. Now the word ‘ambiguous’ is ambiguous itself. According to the *Merriam-Webster Collegiate Dictionary*, it may mean ‘doubtful or uncertain especially from obscurity or indistinctness’, or ‘capable of being understood in two or more possible senses or ways’. If we examine the instances in the poetry of such grand masters of enjambment as Milton, Shelley, Keats, the vast majority are ambiguous in the first sense. They don’t convey two different meanings, but render the gestalts less distinct, more emotional. Critics are, of course, fascinated by those instances that are ambiguous in the second sense. Let me present my position with reference to an intriguing example from a recent paper by Natalie Gerber, ‘Intonation and the Conventions of Free Verse’, where she points out both kinds of ambiguity in James Wright’s poem ‘A Blessing’. She alludes to the first kind of ambiguity in the lines “And the eyes of those two Indian ponies/Darken with kindness’ (3–4), where the segmentation produces more potential affect than if the complete clause were a single line and intonational phrase’ (though she does not explain how does the segmentation produce that greater potential affect). The second kind of ambiguity is conspicuous in the final three lines:

‘Suddenly I realize/That if I stepped out of my body I would break/Into blossom’ (22–24). [...] (B)y placing the line ending between ‘I would break’ and ‘Into blossom,’ Wright interrupts our expectation that we read ‘break into’ as a phrasal verb, with a conventional meaning. Instead we are invited to hear a pause, reading ‘I would break’ and ‘into blossom’ as two intonational phrases and thus entertaining two semantic possibilities at once: that the speaker shatters, a meaning that goes with his stepping out of his body, and that he unexpectedly blossoms.⁶

With reference to the latter ambiguity, my interest lies in the question how should such an enjambment be performed. Suppose we ‘hear a pause, reading ‘I would break’ and ‘into blossom’ as two intonational phrases and thus entertaining two semantic possibilities at once’. What is the phenomenological quality of the sudden shift of mental set. Thus, in such a performance the sudden switch of meanings might have witty overtones. This would be, perhaps, in congruence with some interpretations of the poem. I ask, however, whether one can imagine or secure a performance in which the two meanings blend more smoothly, that would suggest some more earnest attitude, in case someone endorses a different interpretation. The cosy atmosphere of the poem may indeed support such an interpretation. According to the present conception, one could pronounce the enjambment by having recourse to one intonational phrase, with no measurable pause between the two verse lines, suggesting at the same time discontinuity by lengthening the word ‘break’ and overarticulating the stop release of [k]. This example suggests an additional lesson too. Even in such an apparently obvious case it is, eventually, the performer who decides on the intonation contour(s).

The very perception of poetic rhythm crucially depends on the *aural* perception of the verse line as a separate whole. This implies that in order to realise the rhythmical nature of any metre, whether quantitative or syllabo-tonic, one must borrow from Gestalt psychology the assumption that metric organization is a system that determines the character of its parts or, more precisely, a system in which the whole and its parts determine each other’s character (as we have seen, a

⁶ Natalie Gerber, ‘Intonation and the Conventions of Free Verse’, *Style* vol. 49 no. 1, 2015, pp. 9–34, pp. 13–14.

stressed syllable in a strong position may bear a different character at different points of the verse line, and may impose a different character on the line as a whole). The sequence of the verse lines divides the auditory field into larger perceptual units; the line segments divide the line into two equal (or almost equal) parts. One precondition for poetic rhythm to have psychological reality requires the reader to perceive not only a sequence of more or less regularly alternating stressed and unstressed syllables, but also a higher unit amenable to short-term memory, obtaining a larger unit divided by a sequence of smaller units: 'The perception of that which divides is as necessary to the fact of division as is the thing divided' (Chatman, 1965: pp. 23–24). In our case, this minimum requirement for rhythmicality is satisfied by the verse line (the higher unit) and the regularly alternating weak and strong positions (the lower unit). If the higher unit (the line) is strongly established in perception, it tolerates a great number of deviations at the lower levels. Such deviations are experienced as tension—provided that the higher unit eventually emerges as a strong perceptual entity. The delay of expectation and the belief in subsequent fulfilment involved may work only if there is a larger perceptual unit whose emergence can be trusted.

As to whether the relation between rhythmic analysis is descriptive or normative, this may be illuminated by a polemical exchange with Peter Stockwell, who wrote among other things:

I recognise the different genealogy to which Reuven Tsur belongs. Some of his work, for example in aiming for a set of criteria for the 'best' performance of poetry read aloud by actors, would be regarded as inadmissibly prescriptive in my tradition. (Stockwell, 2008: p. 590)

This is a gross misrepresentation of my position, the brutal murder of a straw man. Stricter prescriptive norms prevail in metrics than in any other area of literary research. Far from pursuing a prescriptive agenda, the distinctive feature of my work in prosody is, precisely, an all-out opposition to those strict prescriptive norms. I argue that some of the most musical poets, like Milton and Shelley consistently violated all the criteria for metricalness hitherto proposed. The existence of such lines, however, cannot refute the normative theories in metrics,

precisely because the verse lines that contain those violations are ‘unmetrical’. Rather, proponents of those theories claim that one of the virtues of their theory is that it can distinguish between a metrical and an unmetrical line. I claim (with Wittgenstein) that we draw a boundary for a special purpose. I translated the normative conceptions into descriptive terms: in departures from the paradigmatic line, scales of difficulty, or of ‘unnaturalness’ can be constructed, on which various poets and theorists may draw the utmost boundary of metricalness at different points. Instead of relying on authoritative rules, according to which a stress maximum in a weak position or a polysyllable with its stressed syllable in a weak position render the line unmetrical, I define the boundary in terms of its purpose: to mark the utmost limit of the performer’s willingness or ability to perform the verse line rhythmically. I have given a descriptive definition to ‘rhythmical performance’: one in which the conflicting patterns of language and versification can be perceived at the same time. This limit varies with the performer’s cognitive and vocal skills, as well as aesthetic conception. If the performer of a line with a stress maximum in a weak position cannot render its patterns of language and versification perceptible at the same time, it falls apart; if the performer succeeds, exceptionally high tension is generated.

One final example: In my paper ‘The Structure and Delivery Style of Milton’s Verse’,⁷ I write among other things: ‘As to the first six lines of *Paradise Lost*, the issue at stake is not which one of the many possible performances is the right one, but whether we can secure a performance that may convey its fluid structure’. Later on I write: ‘There are one thousand ways to perform such a line. Opting for the performance suggested here is not meant to disqualify the other nine-hundred-ninety-nine readings. It merely insists that there should be at least one reading that conforms with the foregoing “divergent”, “suspensive”, “fluid”, construal of the passage.’ In contrast to Peter Stockwell, Terry Brogan suggests that my work is an attempt to represent

the concept of Performance at a fundamental level in the theory, in order to

⁷ Reuven Tsur, ‘The Structure and Delivery Style of Milton’s Verse: An Electronic Exercise in Vocal Performance’, *ESC: English Studies in Canada* 33 (2007), pp. 149–168.

achieve greater refinements of explanatory power, i.e. greater ‘delicacy’ of description, by converting Halle and Keyser’s dichotomous metrical determinations (metrical, unmetrical) to a continuum of finer discriminations of metrical complexity.⁸

NG & DNS: ‘Finer discriminations of metrical complexity’: this seems to us an apposite note on which to end. Many thanks for your conversation, Reuven.

When I set out on this interview, I thought I would be expected to concentrate on intonation. But the questions forced me into a wider perspective, and I found myself expounding in a nutshell my whole perception-oriented theory of metre and rhythmical performance. I suppose that is because, when we start to reflect on intonation, we necessarily get drawn into a far broader series of questions and disciplines: especially, for me, instrumental phonetics and Gestalt psychology.

⁸ Terry V. F. Brogan, *English Versification, 1570-1980: A Research Guide* (Baltimore: Johns Hopkins University Press, 1981) p. 312.