Poetry Reading—Rhythmical Performance:
Triple-Encoding, and Voice Quality
Six case studies

REUVEN TSUR
with the participation of Chen Gafni

––––––––––––––––––––––––––

For the past forty years or so I have been exploring the sound structure of poetry and its impact on the perceived effects of poems.[1] My first English publication was on versification (Tsur, 1972). I have elaborated, among others, a "perception-oriented" theory of metre (Tsur, 1977), a theory of rhythmical performance (Tsur, 1977; 2012a) and a theory of musicality and expressiveness of speech sounds (Tsur, 1992). This paper is an adaptation, mutatis mutandis, of an application for a research grant to the Israel Science Foundation. It will bear, therefore, some marks of the conventions of the genre "grant proposals". It presents, very concisely, a few aspects of my work on the vocal performance of poetry.

My basic assumption is that versification and musicality do something to poems that cannot be reduced to meaning or rules. It can only be directly perceived. For the past few decades I have been elaborating a viable metalanguage to talk about such direct perceptions. Much traditional scholarship engages in establishing rules, pointing out the existence of certain sound patterns (sometimes followed by useful statistics) and their relationship to meaning. When there is some deviation from

––––––––––––––––––––––––––

metre, it either tries to "regularize" it, or declares that there arises tension, or points up some relationship to meaning. It rarely asks how the deviant stress or run-on line are perceived as rhythmical. And when it does, it prefers to give abstract rules rather than dwell on what we hear. What I am trying to do, by contrast, is to explore what our ear tells our mind; that is, how the acoustic and phonetic information assumes aesthetic significance. In this I treat the sound dimension of poetry very much like music. I focus on the stream of sounds perceived directly rather than on the referential function of language. Meaning comes in (as we shall see) only after an extended period (in terms of milliseconds) of lingering on the sound information. When there is deviation, I ask in what conditions do we hear, in spite of all, the deviant configuration rhythmical. As a matter of fact, we think we know what "metrical" means, but "rhythmical" is a very fuzzy notion in poetry. Elsewhere I have investigated at considerable length the handling of overtones in music and poetry (Tsur, 2012b: 19–36). Here I will confine myself to some aspects of our current research.

Scientific Background

This project is an instrumental investigation of a theory of rhythmical performance of poetry, originally propounded within the framework of a speculative perception-oriented theory of metre. It has been prompted by such problems as the following. "Iambic pentameter" means that there is a verse unit consisting of an unstressed and a stressed syllable, and that the verse line consists of five such units. In the first 165 verse lines of Paradise Lost, there are two such lines. Why should we speak at all, then, of "iambic pentameter"? Most 20th century prosodic research is a more or less systematic attempt to discover the intuitive rules or criteria followed by the poets. Generative metrics set out to find the rules by which we distinguish a metrical from an unmetrical line. Such a theory cannot be refuted by counterexamples: they are examples of "unmetrical lines". All criteria for metricality hitherto proposed have been violated by the greatest masters of musicality in English poetry. 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 conception requires a systematic theory of "rhythmical performance". Wellek and Warren suggest 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 dimensions: linguistic pattern, versification pattern, and performance. Generative metrists reinvented the first two dimensions and ignore (some argue against) the third one. The present theory is the only one that not only quotes this threefold distinction, but takes it seriously. The instrumental study of poetic rhythm goes back to the early twenties of the twentieth century. One problem with many instrumental studies is that they don't distinguish between the poem's structure and performance: they measure relationships in an accidental performance, mistaking them for relationships in the poem's rhythm. The present theory, by contrast, assumes that a rhythmical performance is a perceptual solution to a perceptual problem posed by the conflicting patterns of language and versification. The present cognitive theory allows to make predictions concerning such rhythmical solutions. Two aspects of short-term memory affect the process: it functions in the acoustic mode, and its span is limited at around 7 monosyllabic words ± 2. That is why the longest verse lines that have no obligatory break are 10-syllable-long. It is assumed that in harmony with the assumptions of the "limited-channel-capacity hypothesis", the performer manipulates his vocal resources in a way that may save mental processing space, so that the verse line may be completed before its beginning fades out in acoustic memory, so that the listener's "metrical set" may reach consciousness, even where its traces in the vocal information are suspended. This is effected by grouping and articulation.

Straightforward instrumental research of poetic rhythm is impossible in principle. The empirical problem is this: there is no way to infer from the machine's output the perceived prominence or the perceived pitch of syllables. 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. Such authorities as D. B. Fry advised

me that I would be better off by relying on my ears rather than on the machine's output. It took me over 25 years to formulate my questions in a way that the machine understands. It is a rather indirect way: apparently trivial findings can be reinterpreted in light of the foregoing theory. The research established a homogeneous set of vocal strategies, by which such diverse metric deviances as enjambment, strings of consecutive stresses and stress maxima in the seventh (weak) position can be rendered rhythmical, in the iambic pentameter. The received view (formulated by Seymour Chatman) was that in case of conflicting intonation contours the vocal performer must choose one contour with its related meanings, and suppress the other. My hypothesis was that conflicting intonation contours *can* be indicated by conflicting vocal devices. In ordinary speech we use redundant phonetic cues; in the rhythmical performance of poetry the same cues may be used in divergence. In the case of enjambment, for instance, experienced performers (poets, leading actors, professors of English and graduate students) tend to be in agreement as to the vocal strategies to be deployed for continuation and discontinuation at the same time. The same cues for continuation and discontinuation are deployed in the rhythmic performance of consecutive stresses. According to the generative metrists, a stress maximum (i.e., a stressed syllable between two unstressed ones) in a weak metric position renders a line "unmetrical". Such poets, however, as Shakespeare, Milton, Keats and Shelley do have recourse to such constructions (more frequently than usually realized), and in a quite systematic manner. Consider Milton's line "Burnt after them to the bottomless pit". This verse line is "unmetrical" under two different criteria of generative metrics proposed by Halle & Keyser, and Paul Kiparsky, respectively. As to the former, the syllable *bot-* is a stress maximum in a weak position; as to the latter, *bottomless* is a polysyllable with its stressed syllable in a weak position. Experienced readers, however, are in agreement as to the performance of such lines. They usually isolate the closed symmetrical group "bottomless pit", over- rather than underemphasize the deviant stress, while also deploying conspicuous cues for continuity between *the* and *bot-*—in accordance with the predictions of the present theory, based on assumptions of gestalt theory and the "limited-channel-capacity hypothesis". These experienced readers are usually surprised when made

aware of what they were doing.

In view of the received views regarding the impossibility to vocally realize conflicting intonation contours, and the impossibility instrumentally to test poetic rhythm as conceived here, it will not be surprising that very little work is done that is comparable to the present work. As to the former issue, in two short articles Katherine Loesch (1965, 1966) spoke of "nondisambiguating intonation contours", for which she was severely rebuked by Seymour Chatman (1966) (and then she became silent). *The New Princeton Encyclopedia of Poetry and Poetics* still endorses Chatman's position. As to the latter issue, Tom Barney in his Lancaster University Masters' Thesis, without being aware of my earlier speculations about conflicting cues, demonstrated instrumentally the conflicting cues for enjambment in Philip Larkin's and John Betjeman's recordings of their own poetry. (This research was instrumental in developing my approach to empirical research.) Most available performance studies are of the kind of my last two issues below: analyzing how intonation and voice quality may contribute to interpretation. One of the best ones is Richard Cauldwell's (1999) study of Philip Larkin's five readings of his "Mr. Bleaney". Five readings enable him to consider alternative solution. But he makes no instrumental analysis, only analyzes intonation with Brazil's notation, and makes general statements on intonation and voice quality.

Though my work is opposed to all the "rule-checking" approaches to poetic prosody, for methodological reasons I state my position mainly with reference to Halle and Keyser's and Paul Kiparsky's varieties of generative metrics.

The case studies below aim to refine my earlier theory published in my 1977 and 1998 [2012a] books. The objectives and expected significance may be illuminated by a polemic exchange with Peter Stockwell (2008), 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 (Tsur, 2008a), would be regarded as inadmissibly prescriptive in my tradition.

This is a gross misrepresentation of my position. 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 (but Shakespeare too) 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. Thus the traditional theories as well as the generative theories are irrefutable. 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 he succeeds, exceptionally high tension is generated.

In my paper 'The Structure and Delivery Style of Milton's Verse' (ESC) 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'.

Owing to recent software developments, I can 'doctor' now recorded readings manipulating the rhythmic solutions they offer; and that's what I am doing in some of my recent publications. It is an empirical way to probe into the question whether a performance can be imagined or secured that allows to perceive the conflicting linguistic and versification patterns at the same time. There are almost insurmountable difficulties in eliciting rhythmicality judgments (Tsur, 2006). But such a doctored performance, not as an imagined performance, can be put—in principle, at least—to empirical test. One could submit a genuine and a doctored version of a line to flesh-and-blood listeners and ask them judge whether any one of them does offer a solution to the conflict of language and versification.

Briefly, nobody knows what rules govern English metre from Chaucer to Yeats, but all agree that they yield admirable results. All the rules proposed by traditional metrists were blatantly violated by the greatest English poets. In my 1977 book I worked out a theory with a radically different conception, including a theory of rhythmical performance. In my 1998 book I provided empirical support to that theory of performance. There is now a second, enlarged edition in 2012. In my current work I aim to refine the empirical research, to ensure consistency and conformity with the fine-grained details.

Case Studies

The ensuing case studies spring from unexpected findings in my work that followed the publication of my 1998 book, and illustrate how novel insights may enrich and further develop the theory presented in that work.

First case study (first issue). 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 I compared two recordings by two leading British actors, Douglas Hodge and Michael Sheen.

(1)   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" [listen]. In Sheen's reading, by contrast' "A" at the beginning of the next line is irritatingly continuous with "express" [listen]. 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 1–2).
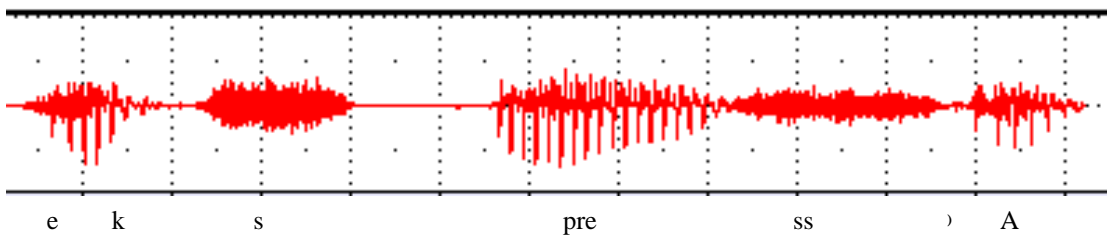

e   k   s   pre   ss   ›   A

Figure 1   Wave plot of "express A" in Hodge's performance ([®] indicates glottal stop)

e   k   s   pre   ss   A
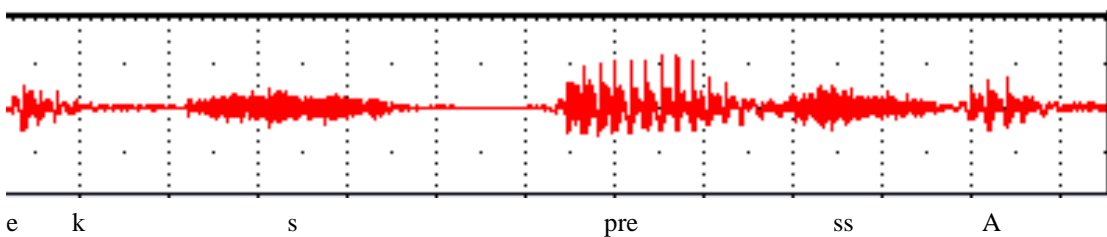
Figure 2   Wave plot of "express A" in Sheen's performance (no glottal stop)

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 [listen]. In addition, Sheen overstresses *our*, violating the iambic cadence of the line. The easiest way to handle such a case is to say that *our* should not bear emphatic stress. My way of handling it is to ask whether a performance can be imagined or secured in which *our* bears emphatic stress, while the iambic cadence is still preserved. 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 [listen]. In future research I will explore the implications of such doctoring exercises.

Second case study (second issue). In poetic structure I have distinguished "divergent" and "convergent" styles. Pope is the archetypal convergent poet, Milton and Shelley the archetypal divergent poets. In my later work I found that in performance too there is a divergent and a convergent delivery style. The performer has considerable freedom in choosing his delivery style. Sometimes the same actor performs the same verse line in a convergent style on one occasion, and in a divergent style on another. I will present my categories of "convergent" and "divergent" delivery styles by way of scrutinizing two recordings by Sir John Gielgud, sixteen years apart, of the last line of Shakespeare's Sonnet 128 (Tsur, 2007):

(2)   To shún the héaven that léads mén to this héll.
         w   s   w   ⌣   s   w   s   w   s   w   s

When I first listened to these two performances by Gielgud, I tried to get an overall intuitive impression of the difference between them, rather than analyze it. I had an unexplained impression that Gielgud 2 is much more complex, artistically

more sophisticated, rhythmically more satisfying. The best way to characterize my impression of Gielgud 1 was, perhaps, by punning on the English idioms "flat-out" and "flat out". The former is usually used as an intensive, that is, a modifier that has little meaning except to intensify the meaning it modifies; the latter suggests "in a blunt and direct manner". Later, when I compared the two readings' handling of the complexities of the verse line, this intuitive contrast was amply accounted for. If you encounter the stretch of language "To shun the heaven that leads men to this hell" in a prose utterance, it may be uttered as a single unit, or will at most be segmented into two segments, the relative clause, and what precedes it. Both readings of Shakespeare's verse are parsed into more segments. Now when you look at the wave plots and pitch plots extracted from the two readings, an immediately-perceived difference becomes conspicuous (Figures 3–4). In Gielgud 1 there is a huge 413-msec pause between "shun" and "the"; and an even longer, 503-msec pause between "heaven" and "that" [listen]. In the wave plot extracted from Gielgud 2, by contrast, no such pauses are visible. Discontinuation is achieved here by means other than straightforward pauses [listen]. Notwithstanding this, one of my associates could hardly believe that there is no pause there in this reading.
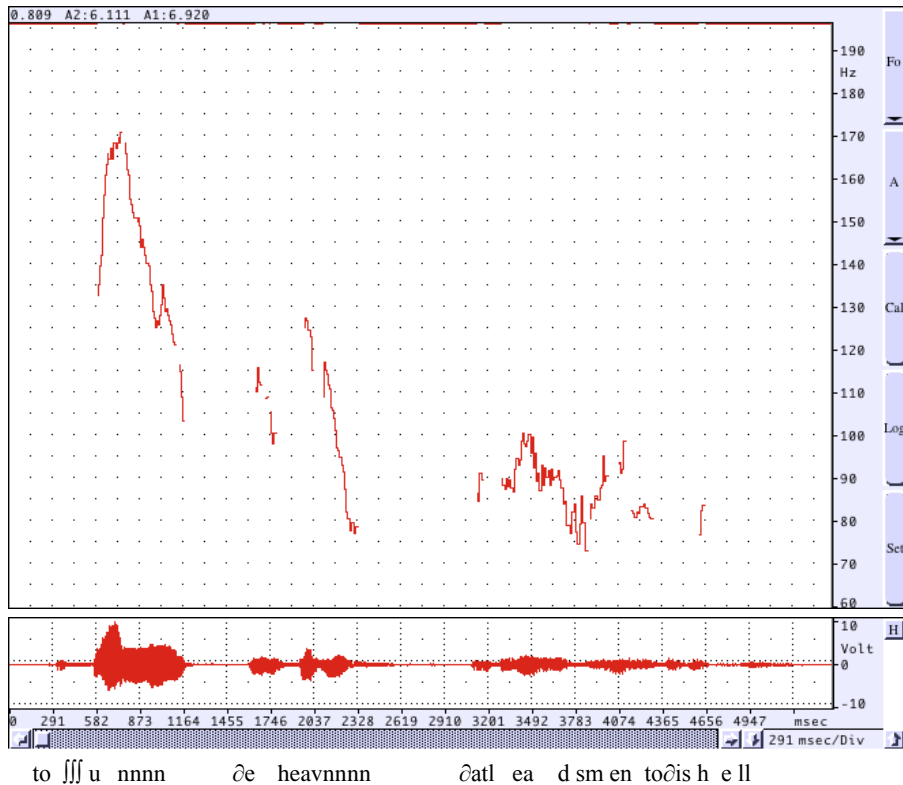
to ʃʃʃ u nnnn   ∂e heavnnnn   ∂atl ea d sm en to∂is h e ll

Figure 3   Wave plot and pitch extract of "To shun the heaven that leads men to this hell" Gielgud 1[2]

What is more, as Figure 4 shows, in Gielgud 2 the words "heaven" and "that" are uttered on one falling intonation contour, effectively grouping "that" backward rather than forward. It is the prolongation and overarticulation of the word-final /n/ that bears all the burden of generating discontinuity at the caesura. Perceptually, what happens here is quite sophisticated. We have got here conflicting cues for continuity and discontinuity. The shared intonation contour and the lack of pause group the word "that" backward; the listener's syntactic knowledge and the sustained /n/ indicating a rest suggest a new start after "heaven". Consequently, a caesura and a "metrical impulsion" across it are perceived at the same time. The line exerts pressure for completion upon which the caesura obtrudes.

---

[2] The lower window presents the wave plot display which shows a plot of the wave amplitude (in volts) on the vertical axis, as a function of time (in milliseconds) on the horizontal axis. The upper window presents a fundamental frequency plot, which displays time on the horizontal axis and the estimated glottal frequency (F_0 = pitch) in Hz on the vertical axis.

Figure 4   Wave plot and pitch extract of "To shun the heaven that leads men to this hell" Gielgud 2

I said above that in Gielgud 1, by contrast, I intuitively felt that the same syntactic juncture was thrust upon the reader "in a blunt and direct manner" and, at the same time, acted as a modifier that has little meaning except to intensify the meaning it modifies. This happens because the grouping cues cluster differently, displaying great redundancy. The beginning of a relative clause constitutes a major syntactic juncture. This is reinforced by an unusually long pause, and is further reinforced by the prolongation of the word-final /n/. It is this overarticulated syntactic juncture that confirms a prosodic event—a caesura. In other words, in Gielgud 1 the cues act in convergence, in Gielgud 2 in divergence. Two differences between the falling intonation contours on "heaven" seem to be quite obvious, one measurable, the other perceptual. First, in Gielgud 1 the falling intonation contour of "heaven" is much longer than in Gielgud 2: it begins at 127.457 Hz (much higher than in the other reading), and falls to 77.915 Hz, slightly below the bottom line of "that" in the other reading. Second, concomitantly, the falling curve

in Gielgud 1 arouses a feeling of "homecoming", whereas in Gielgud 2 there is a feeling that the curve fails to reach the point of rest, demanding completion. What is more, the /n/ is prolonged at this unsatisfactory point, generating a sense of arrest and a sense of impulsion across it at the same time. In Gielgud 1, by contrast, the "homecoming" of the falling intonation contour coincides with a major syntactic juncture, and the beginning of a longish pause, and a longer-than-usual word-final /n/. Exceptionally great stability is achieved. The significance of the duration of a speech sound may be indicated by comparing it to the duration of other speech sounds in the sequence. In Gielgud's readings /n/ is about 1.77 times longer than the combined duration of the preceding sounds h+ɛ+v. By comparison, in the Audio Edition of *Merriam-Webster's Collegiate Dictionary* and in Simon Callow's reading of the same line the combined duration of the preceding sounds is over twice as long as that of /n/. In The Marlowe Society's reading /n/ is insignificantly longer than the combined duration of the preceding sounds. A similar story can be told, mutatis mutandis, about the sequence "shun the". In Gielgud 1 there is a longish pause between them (413 msec); in Gielgud 2 they are run one into the other. In both readings /n/ is considerably longer than the combined duration of the preceding sounds; in Gielgud 2 its relative duration is insignificantly longer than in Gielgud 1. In *Merriam-Webster's Collegiate Dictionary,* by contrast, /n/ is considerably shorter. As after "heaven", after "shun" too Gielgud 1 resorts to redundant cues: the prolongation of /n/ reinforces discontinuity that is also signified by a longish pause; whereas in Gielgud 2 it indicates discontinuity where the two words are run one into another. There are good prosodic and syntactic reasons (caesura and syntactic juncture) to indicate discontinuity after "heaven", with or without a pause. After "shun" it has neither syntactic, nor prosodic justification. It seems to be gratuitous—unless it has some rhetorical or paralinguistic justification. It is, perhaps, prosodic mannerism. In the proposed research I will explore the two delivery styles.

Third case study (third issue). I made an attempt to assess how readers respond to Gielgud's two readings. This led me to realize that there is a crucial problem in such empirical research, which I want to explore. I needed informed informants;

but I also had to request them to try to suspend their professional knowledge. I made Gielgud's readings available on a webpage and asked some of my colleagues to respond to them. I also sent a request to the PSYART and Coglit lists. I received all in all five fairly detailed responses, indicating that the task proved quite difficult. Some of the responses overlapped in some respects, but were at variance in others. The aesthetic event of poetry recital is very complex. The responses I received suggest that in some instances my informants responded to different aspects of the same event. The event may be consistently and quite thoroughly described by the tools offered here. One must, however, realize that in a poll like this all the respondents will respond only to a small subset of aspects. Thus, even widely different responses may be consistent with one another, and relate to the same underlying comprehensive description. These difficulties are amplified by such comments as "I thought the rhythm of number one was subsumed to the acting so that it might have easily gone unnoticed to my unpracticed ear. My attention was called to the rhythm of number two and it was somehow distracting and thereby detracting". I expect this experimentation with listeners' response to pose the worst pitfall in the research.

Fourth case study (fourth issue). I will explore the "triple encodedness" of phonetic cues in metered dramatic speech. The phonetic cues that serve to identify ordinary speech sounds may be manipulated such that they provide information about two additional dimensions of the text: its emotive import and rhythmic organization (Iván Fónagy speaks of the "dual-encodedness" of phonological and emotive information).

> (3)　　Now is the winter of our discontent
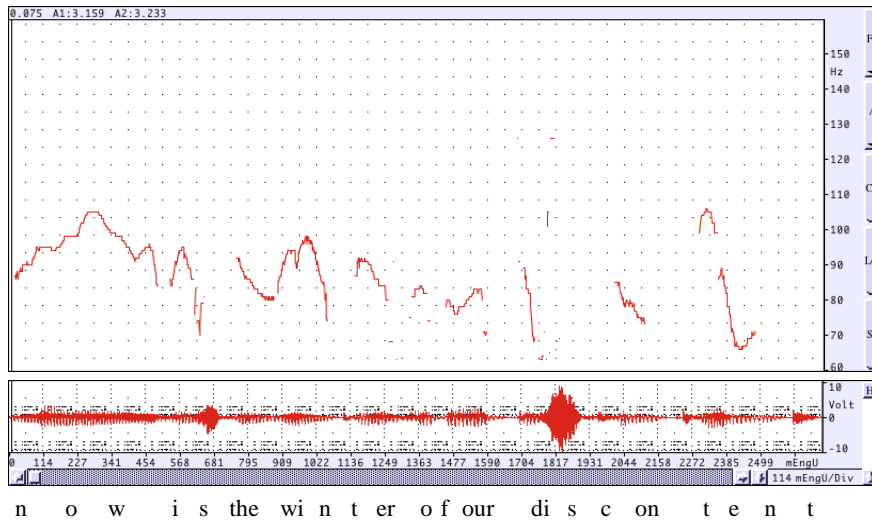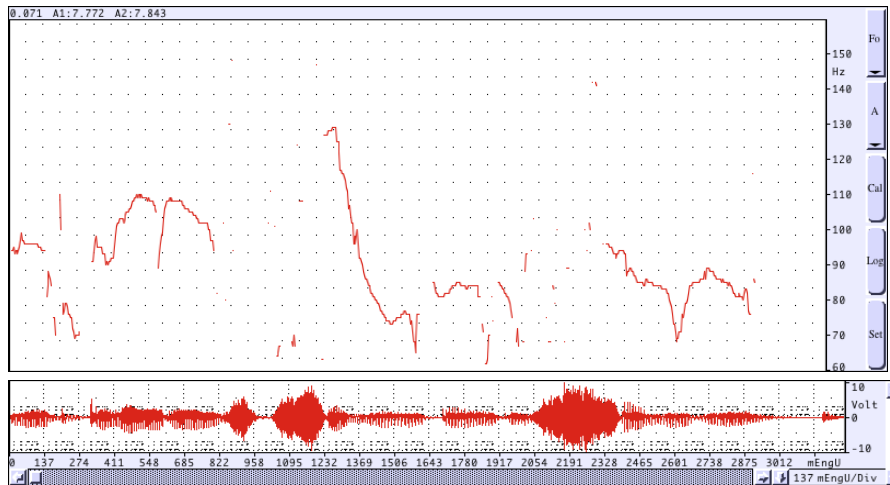> 　　　Made glorious summer by this sun of York

Figure 5   Wave plot and pitch contour of "Now is the winter of our discontent"

I am going to discuss Simon Russel Beale's reading of this excerpt [listen].[3] Figure 7 shows that the line-final *discontent* ends with a stop release. This stop release is exceptionally loud and exceptionally long; and is preceded by a minute pause (28 msec). Compare this to the stop release at the end of *York* in the next line. Here the plosion is much shorter and much weaker. It is, however, preceded by an exceptionally long pause (169 msec—in midword!). In spite of its excessive duration, it is not perceived as a pause, but as an articulatory gesture: extended closure of the vocal track, to overarticulate the [k]. In this instance there appears to be a trade-off between the amplitude and duration of the release and the preceding pause [listen].

---

[3] I have elsewhere discussed since, at considerable length, excerpts from seven recordings of this soliloquy (Tsur, 2012a: 335–373).

ma de gl o ri ou s s umm er by thi s s u n of Yor k

Figure 6 Wave plot and pitch contour of "Made glorious summer by this sun of York"
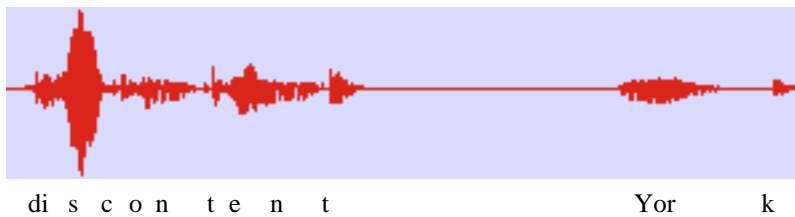
di s c o n te n t Yor k

Figure 7 Wave plots of "discontent" and "York" excised from a reading of excerpt 3.

At the end of the second line, line boundary and sentence boundary coincide; so, there is no syntactic demand here for continuation. Indeed, the phonetic cues are, again, in harmony, all of them indicating discontinuation. This is what I have called redundancy. The final monosyllable, *York*, is the longest syllable in the first two lines (544 msec). The final rising-and-falling intonation curve too may effectively contribute to closure. Considering that there is no prosodic problem here to solve, the line-final stop release with the preceding excessive pause may be judged very much exaggerated. It is here where expressive force and overdetermination come in. The overarticulated line-final stop does not serve merely to clearly articulate a juncture of a line-boundary and sentence-boundary; it serves an expressive function too. The distorted pronunciation of a phoneme may be decoded as a phoneme, as some expressive effect, and as some prosodic effect.

We have just considered the prosodic function of the overarticulated oral stops at the end of the first two lines: to clearly articulate the line boundary. Notice, however, this: a tendency to overarticulate oral stops may be an indication of certain personality traits, such as an assertive, determined, firm attitude. This description fits Gloucester extremely well. I submit that the overarticulation of such word-final stop releases may also be a part of the means by which this particular actor, in this particular performance, characterizes Gloucester as a relentless, determined person. Such an interpretation can rely on many more overarticulated stops in this speech.

Fifth Issue. My views on the tone and mood of the speaking voice have been crucially influenced by Iván Fónagy's "vocal gestures" approach. In the investigation of poetry readings by professional actors we must remember Fónagy's (2001: 88) caveat: "The actor does not express his true emotions: he displays emotive attitudes. The psycho-physiological process of simulated and real emotions probably differ". Both poetic rhythm and emotive attitudes are perceptual qualities. Thus I am exploring vocal correlates which, in turn, may illuminate perceptual qualities. Fónagy discusses at great length (ibid., 87–173) the articulatory and acoustic correlates of the vocal expression of emotions. I submit that in many instances actors have recourse to a subset of these vocal and articulatory devices, sometimes just enough to have the audience perceive the special mood or emotion suggested by the speaking voice. Such a conception is supported by Fónagy's experiments with synthesized speech, in which certain vocal variables believed to be of emotive importance were systematically varied. Listeners "recognize" certain configurations of as few as two acoustic cues as expressing some tone or emotion; and when a small change is made in one of the cues, they "recognize" a different tone or emotion. Thus, for instance, to indicate *threat,* "tonal rise had to be supported by a gradual and considerable rise in intensity; without such a parallel rise of intensity level, the utterance was usually interpreted as a simple warning" (ibid., 104). Or, "a sudden final rise of the F0 curve [...], accompanied by a radical fall in intensity, suggested coquetry in somewhat more than 50% of the cases. The same F0 function failed to elicit any

coquetry votes if the intensity diminished less abruptly. Such a variant suggested only 'interest'" (ibid., 105). This conception must be complemented by Christine Bartels' (1999) work on intonational meaning where, too, subtle and minimal oppositions such as that between a falling boundary tone and a final fall-rise may affect the emotive attitudes of statements, though rarely affecting their truth value. In Fónagy's work there is a huge experimental repertoire of moods and attitudes and their perceived correlation with subtle and minimal vocal cues of the kind of the foregoing examples. Each one of these experimental findings will serve as hypotheses in my work on voice quality. I have already promising beginnings in my book on "Kubla Khan".

Fifth case study (sixth issue)
Overtones act differently in noises and music on the one hand, and in speech on the other. Overtones are higher tones produced *simultaneously* with the fundamental, with which they are "grown together" into a complex, unitary musical note. Owing to different overtone structures we can tell a wooden from a metallic noise, or whether the same note is played on, e.g., a violin or a flute. In speech, vowels are uniquely determined by concentrations of overtones at different pitches (called formants).

   In course of our instrumental study of the rhythmical performance of poetry we noticed that some performers sometimes had recourse to a certain voice quality, deviating from that of the context, for emotive expression. This quality could be roughly described as "softened", "tender". It sometimes assumed an emotional tint of "yearning", sometimes of "affection" or some other tender emotion. Since this voice quality has no name, and we have only begun to identify its distinctive phonetic features, we will provide an ostensive definition of it. Consider the following verse line from the great Hungarian poet's, Mihály Babits's "Esti Kérdés" (Evening Question):

   (4)   merengj a messze multba visszaríván
               muse, weeping and yearning back to the distant past

We have explored on the computer some rhythmic problems of this poem and the rhythmic solutions offered to them by three leading Hungarian actors and the poet himself (recorded in 1936). In one of the readings, by Ferenc Kállai, we found that in contexts of musing or yearning, he sometime deviated from his ordinary voice quality, by somehow softening his voice. In the above verse line he softened his voice when uttering the first word, "merengj" (muse) [listen].

Here apparently we encounter a problem. Every tone has its unique overtone structure, whether played on a violin or a flute. Obviously, the same sounds, with the same overtone structure, can be played on both the violin and the flute. So, how can the ear distinguish between the sounds of the two instruments, based on their overtone structures? The same problem arises, mutatis mutandis, with the changing voice qualities uttering the same words uniquely determined by their formant structure. The answer seems to be that, among other things, the various instruments (or voice qualities) differ in the relative audibility of their overtones.

As we said, every vowel can uniquely be identified by its unique concentrations of overtones called "formants". Normally, spectrograms show up to the first five formants (marked F1, F2, F3, etc.). The first two (lowest) formants are crucial for the identification of a vowel; the rest are more relevant to the sound quality. At the beginning of our project we were looking for a workable hypothesis to begin with, based on the foregoing distinctions. We had reasons to assume that the clue lied in overemphasizing the higher overtones at the expense of the lower ones. So we started with correlations between measurements of the higher overtones and those of the lower ones, and relative to each other. To make a long story short, we have found that the cumulative impact of slight differences in a number of variables is involved in the overall perceived qualities. As to the relative audibility of the various formants, preliminary measurements do suggest that in this and some other readings such a relative overemphasis of the higher formants is the case.[4]

Sixth case study (seventh issue). Intonation and voice quality contribute to

---

[4]   In this section Chen Gafni is co-author.

interpretation:

>   (5)   And all should cry, Beware! Beware!
>   His flashing eyes, his floating hair!

If we compare Roger Lloyd Pack's reading to other ones, his breaks are more frequently indicated by a pause; also, these pauses are considerably longer. As we shall see, in these two lines Pack compensates for the pauses by a sophisticated use of intonation. At any rate, these two short lines (eight-syllable-long each) are divided by syntax into three and two intonation units, respectively. Lloyd Pack introduces an intriguing twist into his text, displaying the voices of two speakers one atop the other. There is the excited voice of "all" who would "hear me"; and the voice of "I" who reports it. Most reciters conflate the two voices into one stream of excited exclamations. The onlookers cry excitedly; and the speaker in the actual situation imitates, as it were, their excited cry. Lloyd Pack conveys the two as different, even though he himself, as a reciter, has only one voice (the same is accomplished by Alex Jennings, with different vocal cues; see Tsur, 2006b). He performs this by having recourse to the boundary intonation known as "fall-rise" [listen].

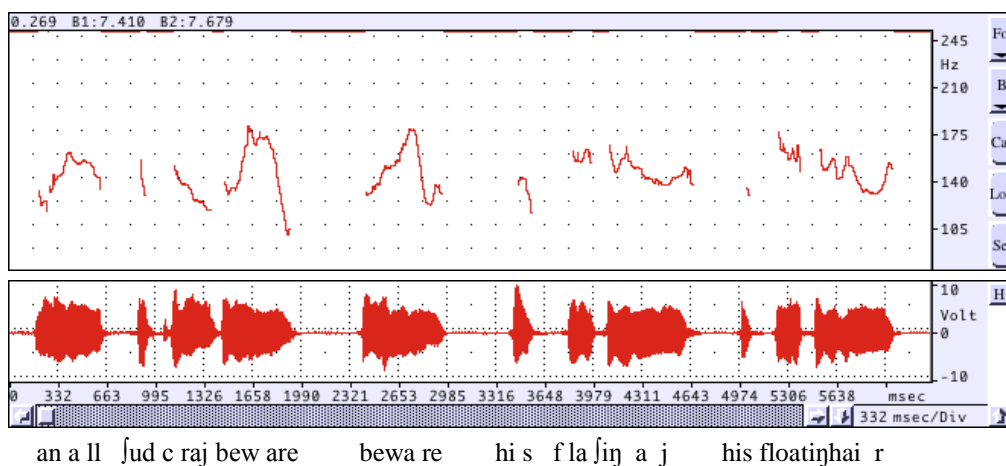an a ll ∫ud c raj bew are   bewa re   hi s ∫ la ∫iŋ a j   his floatiŋhai r

Figure 8 Wave plot and pitch extract of "And all should cry, Beware! Beware! / His flashing eyes, his floating hair!" read by Lloyd Pack.

Adapting to our purpose one of Christine Bartels' observations: "both the [Low]

nuclear accent and the [High] boundary tone are marked options that appear to contribute to the utterance's special connotation" (part of which, in our case, may be an inconclusive, pensive mood), but the basic illocutionary force of this utterance is still that of an assertion, or imposition, or exclamation (1999: 35). The falling intonation curves suggest here a powerful unqualified attitude—warning or horror, or both. It is the attitude of the onlookers in the hypothetical situation reported. The rising boundary tone, by contrast, conveys the unassertive, inconclusive mood of the speaker in the actual situation. The rising boundary tone by itself cannot give more specific information than, e.g., "an unassertive, inconclusive mood"; all the rest must be gathered from the context. Thus, for instance, the speaker in the actual situation may be less than certain that he would ever recover the lost melody, and so have doubts whether those unqualified cries would ever be heard. In the present case one may notice an evasive change of voice quality between the first and second tokens of "Beware!". This change takes place precisely when the speaker shifts from the falling boundary tones at the end of "and all should cry", and the first "Beware" to the fall-rise at the end of the second "Beware", and of each of the two noun phrases "his flashing eye" and "his floating hair".

Works cited

Barney, Tom (1990) "The Forms of Enjambment". University of Lancaster unpublished MA thesis.

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

Cauldwell, Richard (1999) Openings, rhythm and relationships: Philip Larkin reads *Mr. Bleaney. Language and Literature* 8: 35–48.

Brogan, Terry V. F. (1999) *English Versification, 1570 - 1980: A Research Guide. Versification: An Electronic Journal of Literary Prosody.* Available Online: http://www.arsversificandi.net/resources/index.html

Chatman, Seymour, (1966) "On the 'Intonational Fallacy'", *QJS* 52: 283-286,

Fónagy, Iván (1971) "The Functions of Vocal Style", in Seymour Chatman (ed.), *Literary Style: A Symposium.* London: Oxford UP. 159–174.

Fónagy, Iván (2001) *Languages within Language: An Evolutive Approach.* Amsterdam: John Benjamins (Foundations of Semiotics Series).

Loesch, Katherine T. (1965) "Literary Ambiguity and Oral Performance", *QJS* 51: 258-267.

Loesch, Katherine T. (1966) "Reply to Mr. Chatman", *QJS* 52: 286-289.

Preminger, Alex and T. V. F. Brogan (1993) *The New Princeton Encyclopedia of Poetry and Poetics.* Princeton: Princeton UP

Stockwell, Peter (2008) Cartographies of cognitive poetics. *Pragmatics & Cognition* 16:3, 587–598.

Tsur, Reuven (1972) "Articulateness and Requiredness in Iambic Verse". *Style 6:* 123-148.

Tsur, Reuven (1977) *A Perception-Oriented Theory of Metre.* Tel Aviv: The Porter Institute for Poetics and Semiotics.

Tsur, Reuven (1992) *What Makes Sound Patterns Expressive: The Poetic Mode of Speech-Perception.* Durham N. C.: Duke UP.

Tsur, Reuven (1998) *Poetic Rhythm: Structure and Performance—An Empirical Study in Cognitive Poetics.* Bern: Peter Lang.

— (2012a) *Poetic Rhythm Structure and Performance: An Empirical Study in Cognitive*

Poetics. Revised and expanded edition. Brighton and Portland: Sussex Academic Press.

Tsur, Reuven (2006a Winter) "Delivery style and listener response in the rhythmical performance of Shakespeare's sonnets" *College Literature* 33.1 (special issue on "Cognitive Shakespeare: Criticism and Theory in the Age of Neuroscience", guest-edited by Patrick Colm Hogan and Lalita Pandit) pp. 170-196.

Tsur, Reuven (2006b) *"Kubla Khan"—Poetic Structure, Hypnotic Quality and Cognitive Style: A Study in Mental, Vocal, and Critical Performance.* Amsterdam: John Benjamins.

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

Tsur, Reuven (2012b) *Playing by Ear and the Tip of the Tongue: Precategorial information in poetry.* Amsterdam / Philadelphia: John Benjamins.

Willett, S.J., (2001) Review of *Poetic Rhythm: Structure and Performance—An Empirical Study in Cognitive Poetics*, by Reuven Tsur. *Journal of Pragmatics* 33: 333–338.

Willett, S.J. (2004) Reconsidering Reuven Tsur's *Poetic Rhythm: Structure and Performance—An Empirical Study in Cognitive Poetics. Journal of Pragmatics.* 37: 497-503.

Recorded Readings

Beale, Simon Russel *et al.* reading *William Shakespeare: Great Speeches and Soliloquies.* Naxos AudioBooks Na 20 1512.

Gielgud, Sir John reading *Shakespeare: Ages of Man*. HarperCollins AudioBooks HCA 53.

Gielgud, Sir John reading *Sonnets of William Shakespeare*. Caedmon SRS 241 C-D.

Hodge, Douglas reading *John Keats*. Hodder Headline AudioBooks HH 186. (1995).

Kállai, Ferenc reading Babits's "Esti Kérdés", on *Varietas Delectat* 3 (Hungaroton 14279).

Lloyd Pack, Roger reading *S. T. Coleridge*. Rough Winds Productions. (2000)

Sheen, Michael reading *Great Poets of the Romantic Age*. Naxos AudioBooks NA20
    2112. (1994).