

Reuven Tsur

Aesthetic Qualities as Structural Resemblance
Divergence and Perceptual Forces in Poetry

In this paper I will explore two sets of related features of poetic structure, which run all through my work, since my first English publication to my latest articles: convergence-and-divergence, and perceptual forces.¹ I will mention convergence for the sake of comparison only, and concentrate on divergence and perceptual forces.

Traditional poetics has largely described the structures underlying these features, but had to resort to impressionistic means to point out their contribution to emotional qualities. Cognitive poetics, by contrast, is tailor-made to deal with that aspect of poetry in a principled manner. As to aesthetic qualities, the following example may illuminate their nature. The adjective “sad” has different meanings in the sentences “My sister is sad”, and “The music is sad”. In the former, it refers to the mental processes of a flesh-and-blood person. In the latter it does not refer to the mental processes of the sound sequence. Nor does it refer to the mental processes the music arouses in the listener. One can be perfectly consistent when saying “That sad piece of music inspired me with happiness”. It reports, rather, that the listener has detected some structural resemblance between the music and an emotion, such as low energy level, slow movement, and a withdrawn, unassertive attitude suggested by the minor key.² In this sense, “sad” refers to an aesthetic quality of the music. Cognitive Poetics provides a conceptual system that allows to explore similar aesthetic qualities in specific instances of poetry. We will isolate two structural aspects of emotions: relative disorganization and deviation from normal energy level.

How do systems of music-sounds and verbal signs assume perceptual qualities endemic to other systems, such as human emotions or animal calls? At the present stage of my argument I only want to point out that the resources available in the target systems impose severe strictures on the process. Usually only very few

-
1. The sound files for this paper are available online:
http://www.tau.ac.il/~tsurxx/Structural_Resemblance/within_thine_own.html
 2. For a cognitive explanation (in an evolutionary perspective) of the “withdrawn, unassertive” affective character of the minor key see, e.g., Cook and Fujisawa 2006: 9–16; Cook and Hayashi 2008: 318–319. As to the affective character of the major and minor modes, Cook and Hayashi (2008:311) quote Jean-Philippe Rameau, the French composer and author of an influential book on harmony, who wrote in 1722: “‘The major mode is suitable for songs of mirth and rejoicing,’ sometimes ‘tempests and furies,’ [...] as well as ‘grandeur and magnificence.’ The minor mode, on the other hand, is suitable for ‘sweetness or tenderness, complaints, and mournful songs.’”.

features or configurations thereof are available in the target systems that may be shared with the source phenomena. So, the best one can do is to choose the nearest options available in the target system. Minute differences may suffice to transform the perceived character of a complex whole. As Krueger (1968: 100–101) observed, the overall perceived qualities of “total complexes” is determined by minute differences: “It has been observed over and over that the smallest changes in experience are felt emotionally long before the change can be exactly described”.

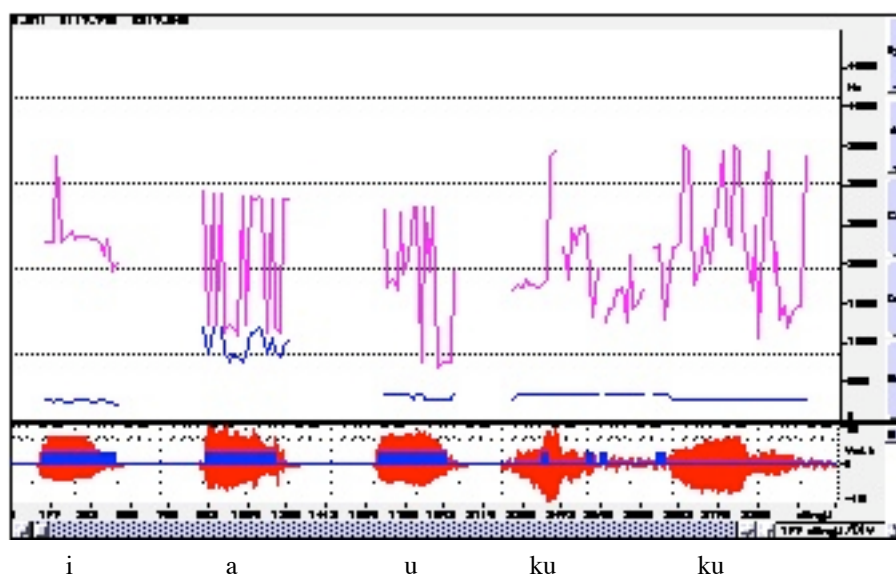


Figure 1 Wave plot, and the first and second formants of the cardinal vowels i-a-u, and of the European cuckoo's call. (Formants are concentrations of overtones that determine vowels and sound colour). Note that the formants of the bird's call are most similar to, but not identical with, the vowel [u] (produced on SoundScope).

In onomatopoeia, the phonological system of a language cannot reproduce the actual sounds of, e.g., the cuckoo's call: neither the minor-third interval, nor the sound quality, nor the abrupt onset. The bird says neither [k] nor [u]. [listen] The only thing one can do is to choose the speech sounds with the nearest formant structure (see figure 1). A symphony orchestra, by contrast, can reproduce the minor-third interval, but not the formant structure of the call. [listen]

kuku

toy symphony

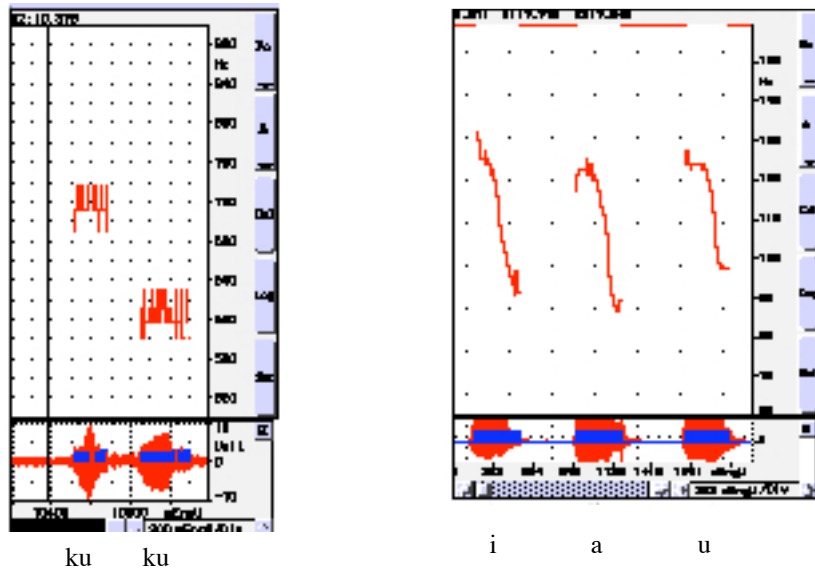


Figure 2 Wave plot and pitch abstract of the European cuckoo's call and of the cardinal vowels read by a professional reader.

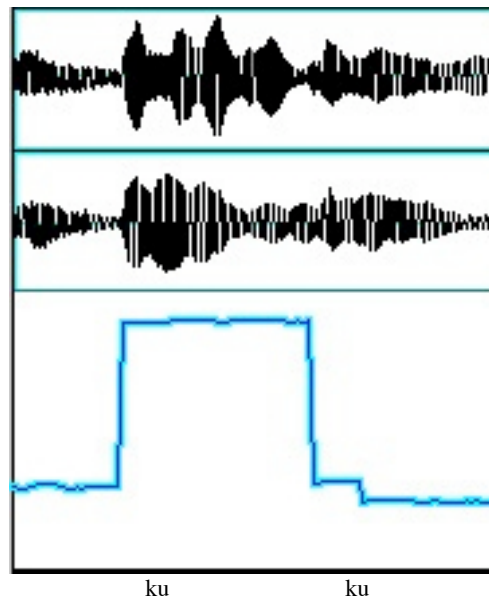


Figure 3 Sound waves and pitch extract of the imitation of the cuckoo's call in L. Mozart's "Toy Symphony" (produced on Praat).

The nearest option to codify in human speech the abrupt onset of the call is the abrupt consonant [k]—all the other features of [k] are irrelevant. In the orchestra, the abrupt onset is indicated more directly (see Figure 3). Thus, the voiceless plosive [k] is a bundle of perceptual features, a subset of which is frequently exploited by the context to suggest some abrupt metallic noises as “ticktack” or “click”; but in the case of “cuckoo” only the perceptual feature [+ABRUPT] is utilized. Thus, the same elements or configurations in a target system may serve as the “nearest option” for a wide range of source phenomena.

In a recent paper I discuss Milton’s verse line

1. And sāt as Prínces, whom the suprême Kíng

w s w s w s w s w s

and quote Milton’s 1809 editor, Henry J. Todd’s comment: “I conceive that Milton also intended the last foot of [this] verse to be a spondee, as more dignified and impressive than the accentuation [...] of *súpreme* on the first syllable” (Todd, 1970: 199). As a side issue, I raise the question what “dignified” may mean with reference to metric structures. It suggests, I claim, that the listener may detect some structural resemblance between the consecutive heavy stresses and the outward manifestations of dignity in humans, such as weightiness, reserve of manner, and clearly-articulated gestures. “Weightiness” in a context of two consecutive stressed syllables encumbered by an unstressed syllable of a disyllable in a strong position suggests “massive”, “hard to deal with” or “demanding great effort”. In a context of dignified human behaviour it suggests “of much importance or consequence”. As to “reserve”, a stressed syllable in a weak position (followed by another stressed syllable) “holds back” the rhythmic movement of the line, whereas a dignified person “holds back”, “controls” his responses, the expression of his emotions or thoughts. As to “clearly articulated gestures” in poetry, I argue in that paper that the rhythmical performance of such constructs as a disyllable with its second, stressed, syllable in a weak position requires exceptionally clear articulation. According to the foregoing analysis, then, within the limited resources of metric structures, consecutive stresses with some additional difficulty are the nearest perceptual options for suggesting a general quality of muchness, slowness and articulateness that can be individuated through the meaning of words as expressing dignity. But notice this. There is no iconic relationship here between metric structure and contents. Rather, the contents individuates as dignified the generalized qualities suggested by the metric structure. Pope’s

2. And ten low words oft creep in one dull line

is not perceived as “dignified”; rather, the two verse lines exploit different potentials of slowness suggested by successive stressed syllables.³

The terms “convergent” and “divergent” are taken from optics where they are applied to rays of light which meet or tend to meet in a focus, and to rays which continually depart from one another. Guilford adopted those terms in the phrases “convergent-thinking” and “divergent-thinking abilities”, referring by them to logical and creative thinking, respectively. One may add that emotional processes are typically more divergent than non-emotional mental processes. This suggests a spectrum: LOGICAL THINKING —► CREATIVE THINKING —► EMOTIONAL PROCESSES. Emotional and non-emotional processes do not constitute a rigid dichotomy, but a continuum. “There is no point on this continuum”, says Elizabeth Duffy (1968: 138), “where a ‘non-emotional’ degree of disorganization of response changes suddenly to an ‘emotional’ degree of disorganization; and there is no point at which a ‘non-emotional’ conscious state changes suddenly to an ‘emotional’ one. These characteristics of experience and behavior show continuous variation rather than separation into hard and fast categories”. The same holds true of the energy-level continuum. I borrowed Guilford’s terms to describe the structural resemblance between certain poetic structures on the one hand, and convergent and divergent mental processes on the other. Now notice this: “disorganization” in divergent poetry is, still, severely constrained by regular meter.

I propose to introduce the distinction “convergence vs. divergence” by comparing two passages in which other things are really equal, in fact, *literally* identical, where the only difference is the line division:

3. But wherefore thou alone? Wherefore with thee
 Came not all Hell broke loose? Is pain to them
 Less pain, less to be fled, or thou than they
 Less hardy to endure? Courageous Chief,
 The first in flight from pain, had’st thou alleg’d

³ As to the desirable potential of slowness see, e.g., Shakespeare’s Sonnet 94:

They that have power to hurt and will do none,
 That do not do the thing they most do show,
 Who, moving others, are themselves as stone,
 Unmoved, cold, and to temptation slow,
 They rightly do inherit heaven’s graces...

To thy deserted host this cause of flight,
 Thou surely had'st not come sole fugitive.
 (*Paradise Lost* IV. 917–923)

4. But wherefore thou alone?
 Wherefore with thee came not all Hell broke loose?
 Is pain to them less pain, less to be fled,
 Or thou than they less hardy to endure?
 Courageous Chief, the first in flight from pain,
 Had'st thou alleg'd to thy deserted host
 This cause of flight, thou surely had'st not come
 Sole fugitive.

Excerpt 3 consists of a series of “straddled lines”. These are sentences run-on from one line to another which themselves, when isolated, form an iambic pentameter line. The run-on lines of Excerpt 3 are rearranged (by James Whaler) into end-stopped lines in Excerpt 4. This rearrangement affects the perceived quality of the passage. Excerpt 3 is perceived as fluid, whereas Excerpt 4 as more stable. When the syntactic unit and the verse line coincide, they reinforce each other’s shape, yielding “strong gestalts”. When the syntactic unit is run on from one line to another, they blur each other, yielding “weak gestalts”. I’ve asked students “Is irony equally subtle in the two passages?” Some students could discern no significant difference. But the rest were in agreement that irony seems to be ‘somehow subtler’ in Excerpt 3. How can we explain this? Semantically and syntactically the two passages are identical.

Gestalt psychologists have produced evidence that strong gestalts are typically perceived as rational, non-emotional, whereas weak gestalts typically display an emotional quality. A similar correlation emerges from findings of the Rorschach inkblot test. In Excerpt 3, the sentences run on from line to line, and the line boundaries intruding upon the sentences blur each other, weakening each other’s shape.

Leonard B. Meyer, who applies gestalt theory to music, accounts for the association of weak and strong gestalts with emotional and intellectual qualities as follows. “Because good shape is intelligible in this sense, it creates a psychological atmosphere of certainty, security, and patent purpose, in which the listener feels a sense

of control and power as well as a sense of specific tendency and definite direction” (Meyer, 1956: 160). Poor shapes generate an opposite atmosphere.

We have noted, however, that the divergent structure in Excerpt 3 seems to affect not only emotional qualities, but irony too, rendering it subtler. Meyer’s formulation may account for this effect too, precisely because it refers to a general psychological atmosphere, rather than a specific attitude.

The ironic attitude typically involves some kind of pretended ignorance. The “psychological atmosphere of patent purpose” inspired by the stronger gestalts in Excerpt 4 subverts, therefore, the tone of elusive ignorance in irony. Weak gestalts, divergent structures, may enhance, then, quite diverse attitudes. Rather than indicating an iconic relationship between form and content, divergent structures generate a “a psychological atmosphere of uncertainty, lack of patent purpose and definite direction”, concreted by various kinds of contents in a variety of more specific attitudes.

Our other term is “perceptual forces”. At the beginning of his book *Art and Visual Perception*, Arnheim demonstrates “the hidden structure of a square” by placing a black cardboard disk in various positions on a white square.

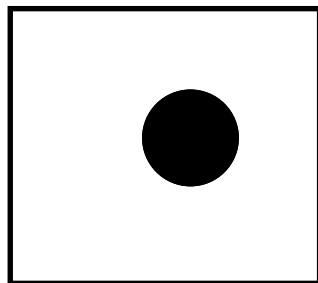


Figure 4

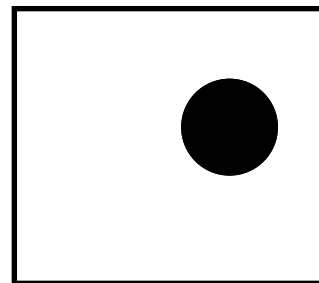


Figure 5

Thus he “maps out” regions of tension and balance. In Figure 4, the disk lies slightly off the centre. “In looking at the disk” he says “we may find that it does not merely occupy a certain place but exhibits restlessness. This restlessness may be experienced as a tendency of the disk to get away from where it is placed or, more specifically, as a pull in a particular direction—for example, toward the center” (Arnheim, 1967: 2).

“Psychologically”, says Arnheim, “the pulls in the disk exist in the experience of any person who looks at it” (ibid., 6). “There is no point in calling these forces

‘illusions’ [he says]. They are no more illusory than colors, which are attributed to the objects themselves, although they are actually nothing but the reactions of the nervous system to light of particular wave lengths” (ibid., 8). “The disk is most stably settled when its center coincides with the center of the square. In Figure 5 it may be seen as drawn toward the contour to the right. With changing distance this effect will weaken or even turn into its opposite” (ibid., 3).

Do perceptual forces exist in verbal structures as well? Fodor and his colleagues used this principle to test the psychological reality of constituent or phrase structure of sentences. The technique is based on the Gestalt assumption that a perceptual unit tends “to preserve its integrity by resisting interruptions” (Fodor and Bever, 1965: 415). In the experiment, subjects listened to a sentence during which a click occurred, and immediately afterward were required to write down the sentence and indicate where the click had occurred. If a phrase is a perceptual unit, subjects should tend to hear a click which occurred during a phrase as having occurred between the phrases. One of their sentences was “That he was happy was evident from the way he smiled”. This sentence has a major break between “happy” and “was”. A click was placed at various positions in this sentence. [...] Each subject heard the sentence with only one click on it. Fodor and Bever found that subjects were most accurate in locating the click which occurred between the two major phrases of the sentence—i.e., between “happy” and “was” in the above example. Clicks occurring before this break tended to be displaced towards the right (i.e., into the break), and those occurring after the break towards the left (i.e., again into the break). A later experiment indicates that even where there can be no acoustic cues for a break, mere syntactic knowledge may evoke such perceptual forces. Fodor et al., however, overlooked one crucial point. If the intrusion occurs in the middle of a perceptual unit, it induces balance and stability; it is only when it occurs between the middle and the boundary that it induces perceptual forces.

For our present purpose, these results have two important implications: first, that perceptual forces do exist in a linguistic environment; second, that perceptual forces in a linguistic environment are crucially influenced by the placement of the intruding event relative to the boundary of the perceptual unit. In poetic prosody there is a further complication. One cannot elicit perceptual forces with the help of some extra-linguistic click. However, syntactic boundaries may intrude upon verse lines perceived as *wholes*; and line boundaries may intrude upon syntactic units. Here the exponents of both the intruding and the disrupted events are conveyed by the same noises, by the *same* words.

In a verse line, a syntactic break at the caesura reinforces balance and symmetry; the nearer to the end of the line, the more it presses toward the end for completion. Consequently, our relief will be greater when the missing part is supplied. This may generate, in certain circumstances, a sharp, witty effect, turning the last string of syllables into a “punch-phrase”, so to speak. In Excerpt 5, from Pope’s *An Essay on Criticism*, there is little that can account for the wit of the second line, except the requiredness of the last word:

5. Some foreign writers, some our own despise,
 The ancients only, or the moderns, prize.

This is a characteristic feature of Pope’s wit. In divergent poetry, the effect is much more sophisticated. A syntactic break near the line boundary exerts pressure toward the boundary; but the verse line is not end-stopped: the sentence runs on to the next line. In such cases a sense of sweeping movement may be generated. Consider, for instance, the following excerpt from Milton’s “On his blindness”, and note the placement of the two tokens of “best”.

6. “God doth not need
 Either man’s work, nor his own gifts. Who best
 Bear his mild yoke, they serve him best. His state
 Is kingly: thousands at his bidding speed,
 And post o’er land and ocean without rest;
 They also serve who only stand and wait.”

The sentence “Who best bear his mild yoke, they serve him best” could constitute an iambic pentameter line, and the repeated “best” would enhance the symmetry and stability of its segments. Had Milton divided this run-on line into 6+4 syllables,

7. Who best bear his mild yoke,
 They serve him best.,

he would have generated a relatively mild divergent movement. As it stands, straddled over two lines, beginning in the ninth position and ending in the eighth position of the next line, the repeated pair of words introduces asymmetry and great instability into the sequence. The nearer an intruding break to the middle of a

perceptual unit, the more it enhances symmetry; the nearer to its boundary, the more it enhances asymmetry and instability. Here the straddled line begins near the line boundary and ends just before the next line boundary. The first token of “best” occurs at the line boundary which, in turn, intrudes upon the complex sentence near its beginning; the second token of “best” occurs at the sentence boundary which, in turn, intrudes upon the line near its end.

Now compare this to the following excerpt from Coleridge’s “The Rime of the Ancient Mariner”;

8. Farewell, farewell! but this I tell
 To thee, thou Wedding-Guest!
 He prayeth well, who loveth well
 Both man and bird and beast.
- He prayeth best, who loveth best
 All things both great and small;
 For the dear God who loveth us,
 He made and loveth all.

While in Milton’s sonnet the two tokens of “best” disturb the balance and induce fluidity, Coleridge’s repeated “well” and “best” generate stability: their two tokens occur at the precise middle and end of the line, generating a sharp, epigrammatic quality. The two syntactic units (ending with “well” or “best”) converge with the two half-lines; in Milton, they diverge. Furthermore, while in Coleridge the relative clause *follows* the main clause: “He prayeth best, who loveth best”, Milton inverts this order, so as to increase the predictive load of syntax, generating suspense: “Who best bear his mild yoke, they serve him best”. Milton’s poem conveys a theologically-laden inner struggle. The theological *ideas* assume an emotional character, the sententious tone of the dictum becomes affectionate, owing to the highly divergent structure of the text, suffused with impetuous perceptual forces, generating a perceptual quality that bears a structural resemblance to powerful emotions. The adjective “mild” softens the utterance both by its meaning, and by blurring the iambic metre.

We have contrasted Milton’s divergent sentence-and-versification structure to a similar but convergent structure in “The Ancient Mariner”. We may, however, contrast it more immediately to the structure of the last line in the same poem:

“They also serve who only stand and wait”. Here the elements of language and versification act in consonance to generate an atmosphere of stability: as in Coleridge, the relative clause comes last; this is the only case in this divergent poem in which a whole sentence entirely converges with the line; and one of the exceptional cases in Milton in which stressed syllables occur only in strong positions, and in all strong positions. Thus, the juxtaposition of diametrically different configurations of language and versification points up the contrast between them, generating a powerful sense of fluidity leading up to an intense sense of stable closure.

We have discussed perceptual forces at some length, in visual perception, the psycholinguistic laboratory, and enjambment in poetry. In music Cooper and Meyer pointed out similar perceptual forces: a steeply rising pitch sequence or intensity sequence (crescendo) has a marked forward grouping effect (it leads, so to speak, forward). The same phenomenon we find in speech, in rising intonation contours. But the perceptual forces can be demonstrated at the sub-phonemic level too, in the alignment of intonation and syllable crest. At this level, the intruding event, the peak of the pitch contour, normally occurs in the middle of the syllabic crest, generating stability; in some instances, however, it occurs late in the vowel, or even on a sonorant after it; and sometimes it occurs earlier than the middle. I have found in my corpus of poetry-readings that late peaking generates an impetuous forward drive; in fact, the later the peaking, the more impetuous is the forward drive. An early peak effects backward grouping and stability. Let us observe peak delay in three recordings of line 7 from Shakespeare’s second Sonnet :

9. To say within thine own déep-súnken éyes

Let us listen to Marlowe Society’s reading of this line. [listen] We will focus here only on certain aspects of the words “say”, “within”, and “own”. The second syllable of “within” (being part of a function word) is perceived as unduly prominent, cued by stress, rising intonation and late peaking. Apparently, metric regularity in the first six syllables “To say within thine own” is straightforward enough, and there would appear to be no reason for such extravagant devices. But the reciter has a real problem here: the caesura occurs in the middle of a prepositional phrase. One must indicate an intruding event after “within”, without disrupting the phrase.

Marlowe Society

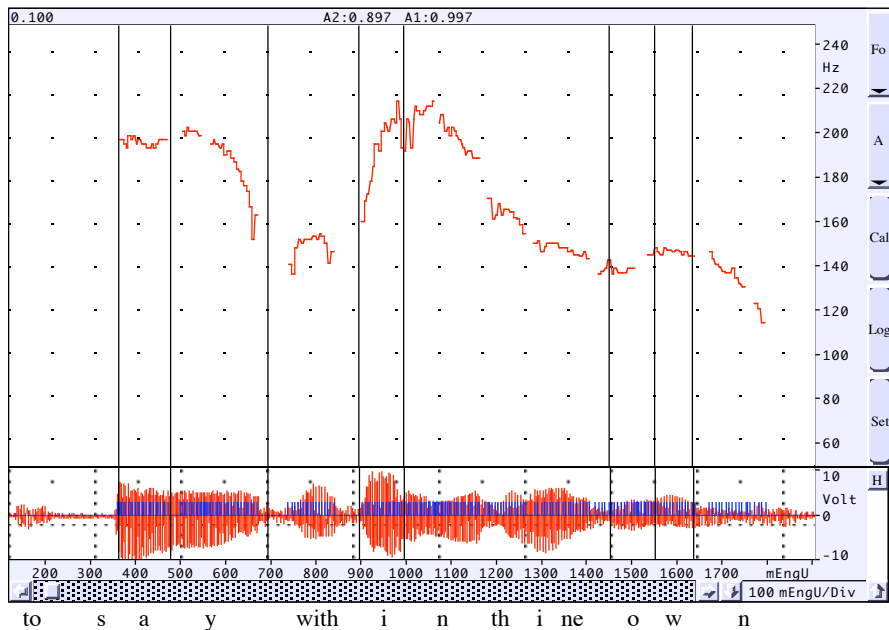


Figure 6 Wave plot and pitch extract of “to say within thine own” in the Marlowe Society’s reading. The markers indicate diphthong and vowel boundaries. Notice the late peaks on *say*, *-thin* and *own*.

The over-articulation of “within”, coupled with an undue stress, without a pause, seems to have here one purpose: to indicate a caesura, without stopping. Listening again to the line confirms that after both words (“within” and “say”) there is an impetuous drive across a discontinuity. [listen again] In the case of “say” it is the contents that compels the reciter to separate the reporting phrase from the reported speech; but if he wants to preserve the line’s perceptual integrity, he must preserve the first four syllables, up to the caesura, as one unit. There is no pause after “say” or after “within”. In both instances, discontinuity is indicated by an exceptionally long word-final sonorant, /j/ and /n/ respectively, and a drastic change of the direction of the pitch contour. In “say” there is a late peak on the second sound of the diphthong. In the second syllable of “within” there is a double peak, one occurring at the end of /i/, the other on /n/. This late peak bestows extreme prominence on “within”, emphasizing the strong position, but also displays an impetuous forward push. On “own” in position 6 there is an additional late peak: it serves the need to group the word forward with the next two words, so as to begin the string of stressed syllables in a strong position.

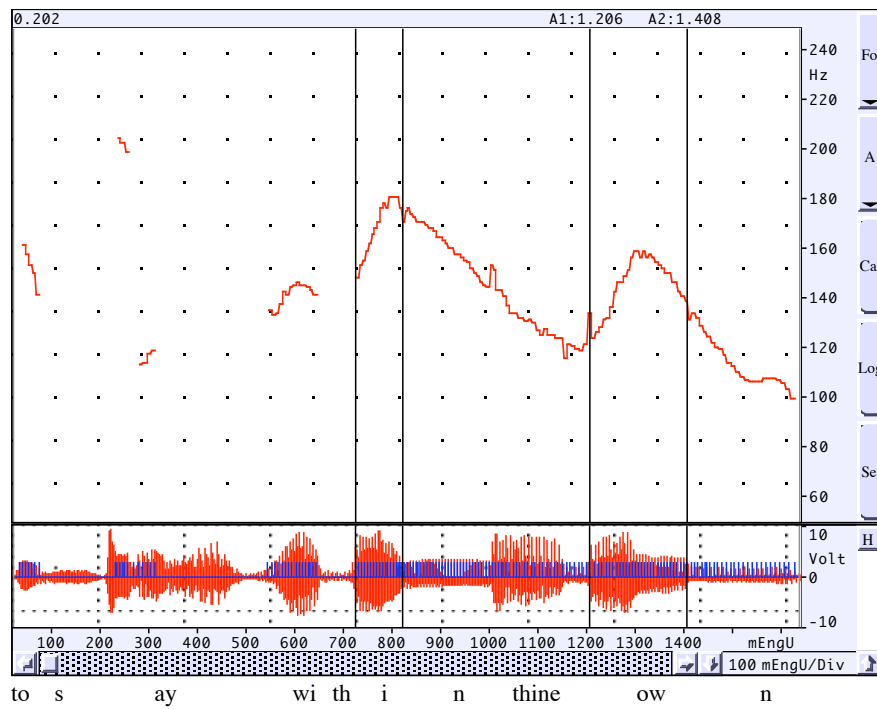


Figure 7 Wave plot and F_0 extract of “to say within thine own” in Callow’s reading. The pairs of markers indicate vowel boundaries. Notice the late peak on *-thin*.

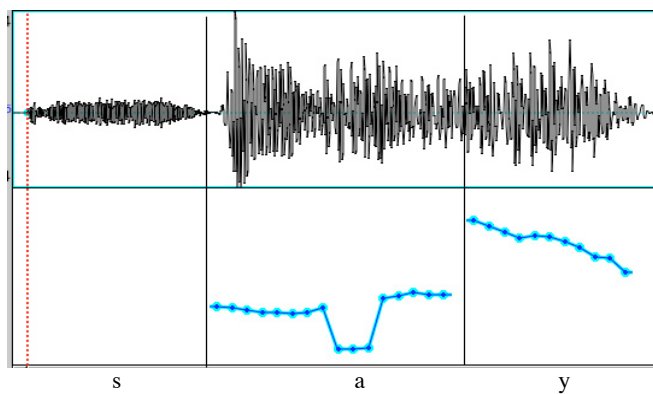


Figure 8 Wave plot and pitch extract of “say” in Callow’s reading (produced on Praat)

Admittedly, the pitch movement the Marlowe Society assigns to the sequence “to say within” is exceptional: these pitch movements cannot be predicted from spoken English prosody, or from any possible metric deviation; it is, indeed justified solely

by the evasive problem of caesura and line integrity. Notwithstanding this, Callow has recourse to similar intonation contours. [listen] He seems to identify exactly the same problems in the first six syllables of the line; and offers exactly the same kind of solution. Here too there is a conspicuous late peak on the second syllable of “within”, and the word-final /n/ is exceptionally long. Figure 8 reveals a rather high and exceptionally late peak on “say”. [listen again]

Callow

Callow

Listening to Gielgud’s performance indicates a conception that is rather similar to that of the other two reciters, but with considerably different emphases. [listen] Pitch resets very high on “say”, with a late peak on /j/ of the diphthong, imposing a forward impetus to the whole line. This drives across an enhanced break indicated both by a pitch discontinuity and a straightforward 100-msec pause.

Gielgud

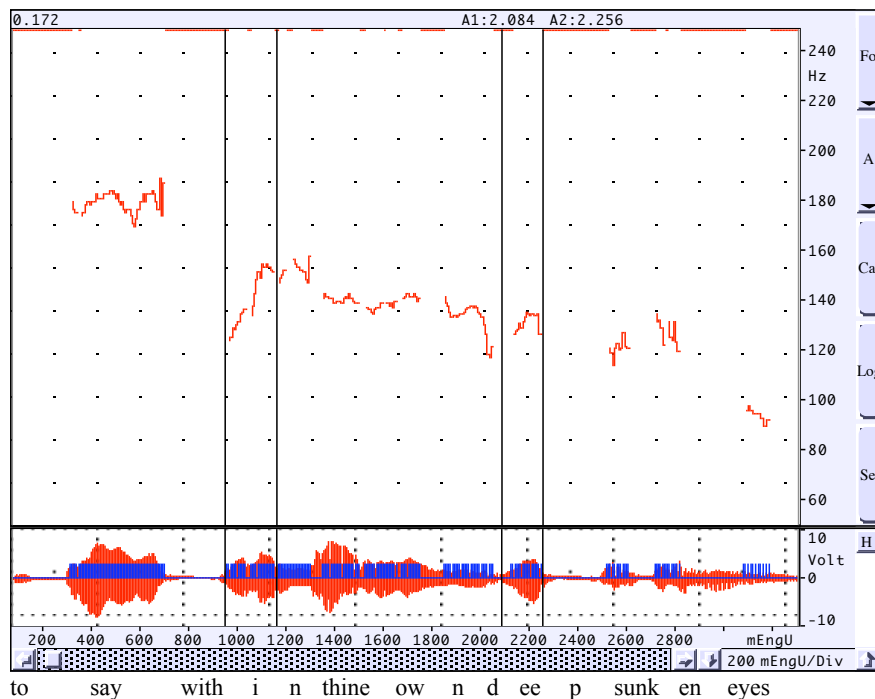


Figure 6 Wave plot and F_0 extract of “to say within thine own deep-sunken eyes” in Gielgud’s reading. The pairs of markers indicate vowel boundaries. Notice the late peaks on *-thin* and *deep*.

In Gielgud’s performance, the peak on the stressed /i/ of “within” is rather moderate; but it occurs very late in the vowel, and there is an additional peak on /n/. The two syllables /(ə)in/ and /(ə)ajn/ belong to two consecutive function words.

They happen to be very similar, but of different duration: the sequence /-(ð)in/ is 239 msec long, of which /n/ takes 148 msec; whereas /(ð)ajn/ is only 155 msec long, of which /n/ takes 67 msec. This relatively long duration contributes to the perception of great stress on “within”, also signaling conspicuous discontinuity. This combination of cues indicates a caesura after “within” and, at the same time, an impetuous drive across it. A similar story can be told of “own” in the sixth position.

[listen again]

Gielgud

Late peaking is a rare, relatively little-understood phenomenon. Gerry Knowles notes that in ordinary speech it usually occurs in the middle of tone groups after a pause. As to its function, Robert Ladd says: “peak delay is said to signal that the utterance is in some way very significant or non-routine” (Ladd, 1996: 99). The more remarkable it is that in this line three leading British actors have recourse to it several times at the same places in the line, most notably on a preposition, at places where no pauses precede them, but some forward thrust is called for, for rhythmic reasons. These actors utilize, then, a kind of vocal manipulation available in language for semantic emphasis—to cope with rhythmic complexities.

I have proposed an approach to emotional qualities in poetry that closely resembles iconicity. It does not, however, pursue an iconic relationship between form and content. The form-and-content approach allows the critic to handle only those instances in which the similarity between form and content exists, or else compels him to read the similarity into them. The present approach replaces this dichotomy by the materials-and-structures dichotomy proposed by Wellek and Warren. It regards both the contents and the formal elements of versification as aesthetically neutral *materials* that can be combined into aesthetic structures. According to the Wellek and Warren model a wide range of elements (which are independent variables) may occur in any combination, and thus the tools offered here may serve to describe any unforeseen combination of elements in a poem. Unforeseen combinations may display unforeseen gestalt qualities, and cognitive poetics may systematically account for them. When we say “The music is sad”, we refer to an aesthetic quality of the music. When we say “The poem is sad”, we may refer *either* to the mere contents of the poem, *or* to an aesthetic quality arising from a configuration of divergent structure, low energy level, slow motion, sad contents.

Contents, “projected world”, word meanings, phonetic structure, metaphor, meter, rhyme, alliteration, are all materials. Structures are their various combinations. Poetic effects arise from the subtle interaction of a great variety of materials. The sequence of stressed and unstressed syllables may converge with or

diverge from the sequence of strong and weak metric positions; syntactic units may coincide with verse lines, or may run on from one line to another; alliteration may work in conjunction with, or against, meter; and so forth. Briefly, they may act in convergence reinforcing each other, yielding exceptionally strong gestalts, sometimes with a pervasive witty quality, sometimes suggesting simplified mastery of reality, as in nursery rhymes. Or they may act in divergence blurring each other, so as to yield an exceptionally weak gestalt with a pervasive emotional or subtle ironic quality. Such divergence may be reinforced by abstract nouns in a landscape defined here and now. “Hypnotic” poetry typically involves exceptionally regular meter–stress mappings, end-stopped lines but unpredictable groupings of lines, alliterations that work both in conjunction with and against meter, frequent repetition of key phrases, high energy level, the irruption of the irrational in the world stratum; and so forth.

Finally, the various configurations need not necessarily comprise homogeneous elements. Consider again “Who best bear his mild yoke, they serve him best”. The strained enjambment, the relative clause preceding the main clause, the consecutive stressed syllables “mild yoke” and “best bear” with the alliteration in adjacent stressed syllables, induce a sense of fluidity and uncertainty. At the same time, the superlative “best”, its symmetrical repetition at the two extremes, and the epigrammatic formulation, all inspire an atmosphere of certainty and stability. Such opposing effects need not generate a conflict between fluidity and stability or mitigate each other. On the contrary, where the powerful drift is established as dominant, the robust stabilizing elements may be perceived as vigorous intrusions at the “wrong” places, enhancing fluidity rather than stability. This divergent, fluid structure has, in Coleridge’s terms, the unpredictability of life and nature, of a “feeling profound and vehement” which, at the same time, is brought “under the irremissive, though gentle and unnoticed, control of will and understanding”—effected by the closure, after the event. But notice this: we don’t perceive an *analogy* between the verbal structure and emotional processes; rather, we perceive the verbal structure *as* displaying an emotional quality.

There are no rules to infer the aesthetic qualities emerging from configurations of aesthetically neutral elements. According to Frank Sibley, we decide that a piece of music is sad by listening or that a piece of poetry is dignified by reading, just as we decide that the book is red by looking or that the tea is sweet by tasting. According to the present conception, disagreement whether a piece of convergent poetry is hypnotic, witty, playful, monotonous, cheerful or suggests simplified mastery of

reality may be due to different mental organizations of the same aesthetically neutral elements.

END OF PAPER

Appendix Observations on Larsen's criticism of the click experiment

Larsen (1971) criticised Fodor & Bever's and related experiments, pointing out that due to insufficient precision of the postulate, results obtained by the click technique seem uninterpretable at present. He pointed out two flaws in their procedure. First, two crucial categories of responses were excluded from their data analyses: responses to clicks objectively located in the major syntactic break itself; and correct subjective locations of clicks objectively located outside the major break. Second, the notion of "constituent" is rather fuzzy. Every word boundary is at the same time a constituent boundary—though at different levels in the hierarchy of constituents. The simplest prediction, therefore, on the assumption that perceptual units resist interruption would be: Noise heard during speech should tend to be located perceptually between words rather than within words. Larsen's experiment refutes this hypothesis.

I agree that the exclusions mentioned by Larsen are hard to understand. I also agree that the notion of "constituent" is rather fuzzy (I have written on this, back in 1973). But it is clear from Fodor et al.'s examples that they are not dealing with just any syntactic constituent, but with constituents delimited by major syntactic boundaries. Thus, for instance, Garret, Bever and Fodor (1966) recorded pairs of sentences such as:

- (1) As a result of their invention's *influence the company was given an*
award. * *
- (2) The chairman whose methods still *influence the company was*
given an award. * *

When subjects were asked where they hear the longest pause in the sentences, they report—as one might expect—that they hear a pause in (1) between "influence" and "the", and in (2) between "company" and "was". The perceived pause thus corresponds to the major constituent boundaries in the two sentences.

Then, the two italicized segments were interchanged. "*Subjects' perception of pause location, however, was unchanged.*" The same was true of click displacement.

As indicated by asterisks in the two sentences above, a click occurred either during “company” or “was”. The perception of click location, however, was significantly different for the two sentences. The click in sentence (1) tended to be heard between “influence” and “the”, and in (2) between “company” and “was”. But remember the sentences were acoustically identical” (Slobin, 1971: 25-26; italics in original). At any rate, click displacement occurs not even toward the boundary between NP and VP, but toward “major syntactic boundaries”—even when the same acoustically identical words are used.

According to the predictions of gestalt theory, the stronger a boundary, the greater its effect on perceptual phenomena. Thus, we shouldn’t expect word boundaries and clause boundaries have the same effect on an intruding click, even though both “words” and “clauses” can be labeled “constituents”.

Anyway, Larsen’s criticism does not affect my argument for two reasons. First, both Fodor et al. and Larsen overlook a crucial point. Perceptual forces are expected to be caused by intrusion not anywhere inside a unit, but between the middle and the boundary of a unit. Intrusion in the middle is expected to reinforce rather than upset stability. Thus, Larsen’s finding that clicks tend to be perceived in midword rather than between words does not necessarily jeopardize my position. Second, my main concern here is not the psychological reality of linguistic units, but two more immediate phenomena: whether perceptual forces do exist in a linguistic environment; and whether they tend to push toward the nearest boundary when dislocated from the middle. In this respect, Larsen’s experiment replicated Fodor et al.’s findings. Larsen concludes that in view of his findings two alternatives and less arbitrary decisions are possible: (1) to reject the hypothesis that linguistic constituents are perceptual units; or (2) to reject the assumption that perceptual units resist interruptions. In view of the importance attributed here to the point in the middle, a third possibility is well worth experimenting too: that an intruding event reinforces perceptual units in their middle, but generate a “push” when they occur between their middle and their boundaries. Thus, for instance, one may expect that a major syntactic boundary reinforces the stability of a verse line when it falls at the caesura, but generates fluidity when it occurs between the caesura and the line boundary. This appears to apply to the alignment of intonation peak and syllabic crest as well. On one point in this respect all researchers appear to agree: that an intonation peak hitting the middle of a syllabic crest is the unmarked option, whereas late and early peaking are marked.

Moreover, I have predicted and empirically found that if you add two syllables at the end of an iambic pentameter line with a caesura after the fourth position, there is a strong tendency to move the caesura to the right, after the sixth position—if the prevailing conditions permit (and sometimes even if they don't). Consider:

10. Invoke thy aid to my advent'rous song.
(*Paradise Lost* I. 13).

In this line, the caesura occurs after *aid* in position 4. Suppose however, that we add two more syllables to the verse line, turning it into an iambic hexameter, thus:

11. Invoke thy aid to my advent'rous song of praise.

If one continues to observe a caesura after *aid*, excerpt 11 is liable to fall apart. Here the caesura, in harmony with the perceptual needs of the iambic hexameter, is automatically shifted to after *my* in position 6, even though this happens in mid-phrase. Here a feeling is generated of a caesura as well as a “sense of impulsion across the [non-existent] break”.

At any rate, Larsen gives no information where exactly in mid-word do the dislocated clicks occur in his experiment; and neither Larsen, nor Fodor et al. indicate whether correct subjective locations of clicks do or do not occur at the middle of a larger syntactic constituent (e.g., clause). This problem is aggravated by another, methodological difficulty: while it is relatively easy to determine where is the boundary of a clause or the middle of a word or a verse line, we have no criteria to determine where is the perceptual middle of a clause. Thus, neither Fodor et al.'s nor Larsen's click experiments may explain why some clicks are correctly located in mid-clause and some not, nor may, perhaps, prove that syntactic constituents have psychological reality; but they strongly suggest that when clicks are subjectively displaced in mid-clause, they move in the expected direction, and seem to corroborate my assumptions. I said "seems to", because to get more compelling results one should re-devise the whole click experiment in view of the importance attributed to the middle point.

References

- Arnheim, Rudolf (1957) *Art and Visual Perception*. London: Faber & Faber.
- Coleridge, Samuel Taylor (1951) *Biographia Literaria*. In Donald A. Stauffer (ed.) *Selected Poetry and Prose*. New York: The Modern Library, 109–428.
- Cook, Norman D. and Takashi X. Fujisawa (2006) “The Psychophysics of Harmony Perception: Harmony Is a Three-Tone Phenomenon”. *Empirical Musicology Review* Vol. 1, No. 2: 1–21.
- Cook, Norman D. and Takefumi Hayashi (2008) “The Psychoacoustics of Harmony Perception” *American Scientist* 96: 311–319.
- Cooper, C. W. and L. B. Meyer (1960) *The Rhythmic Structure of Music*. Chicago: Chicago UP.
- Duffy, Elizabeth (1968 [1941¹]) “An Explanation of Emotional Phenomena without the Use of the Concept Emotion”, in M. B. Arnold (ed.), *The Nature of Emotion*. Harmondsworth: Penguin. 129–140.
- Fodor, J., A. and T. Bever (1965) “The Psychological Reality of Linguistic Segments”. *Journal of Verbal Learning and Verbal Behavior* 4: 414–420.
- Garrett, M., T. Bever and J. A. Fodor (1966) “The Active Use of Grammar in Speech Perception”. *Perception and Psychophysics* 1: 30–32.
- Guilford, J. P. (1970 [1959¹]) “Traits of Creativity”, in P. E. Vernon (ed.), *Creativity*. Harmondsworth: Penguin. 167–188.
- Knowles, Gerry 1992. “Pitch Contours and Tones in the Lancaster/IBM Spoken English Corpus”, in Gerhard Leitner (ed.), *New Directions in English Language Corpora — Methodology, Results, Software Developments*. Berlin: Mouton de Gruyter. 289–299.
- Krueger, F. (1968 [1928¹]) “The Essence of Feeling”, in Magda B. Arnold (ed.), *The Nature of Emotion*. Harmondsworth: Penguin. 97–108.
- Ladd, Robert D. (1996) *Intonational Phonology*. Cambridge University Press.
- Steen F. Larsen (1971) “The Psychological Reality of Linguistic Segments Reconsidered”. *Scand.I. Psychol.*, Vol. 12: 113–118.
- Meyer, Leonard B. (1956) *Emotion and Meaning in Music*. Chicago: Chicago UP.
- Sibley, Frank (1962) “Aesthetic Qualities”, In Joseph Margolis (ed.), *Philosophy Looks at the Arts: Contemporary Readings in Aesthetics*. New York: Scribner. 63–88.
- Slobin, Dan I. (1971) *Psycholinguistics*. Glenview Ill. and Brighton, England: Scott, Foreman & Co.
- Todd, Henry J. quoted by Youmans, Gilbert (1989) “Milton’s Metre”, in *Phonetics and Phonology*, Volume 1: *Rhythm and Meter*. New York: Academic Press. 341–379.

- Wellek, René & Austin Warren (1956) *Theory of Literature*. New York: Harcourt, Brace & Co.
- Whaler, James, 1956, *Counterpoint and Symbol: An Inquiry into the Rhythm of Milton's Epic Style*. Copenhagen: Rosenkilde (= Anglistica 6).

Recorded Readings

- Callow, Simon reading *Shakespeare's Sonnets*. Hodder Headline AudioBooks HH 185.
- Gielgud, Sir John reading *Sonnets of William Shakespeare*. Caedmon SRS 241 C-D.
- The Marlowe Society and Professional Players reading *Shakespeare: The Sonnets*. Argo ZPR 254.

Audio processors

- SoundScope 16/3.0 (ppd)
- Praat 5.0.43

*

“Jean-Philippe Rameau, the French composer and author of an influential book on harmony, wrote in 1722: ‘The major mode is suitable for songs of mirth and rejoicing,’ sometimes ‘tempests and furies,’ [...] as well as ‘grandeur and magnificence.’ The minor mode, on the other hand, is suitable for ‘sweetness or tenderness, complaints, and mournful songs.’” (Cook and Hayashi 2008:311).