Keynote lecture at the Symposium
"Cognition and Poetics: Minding Language and Literature"
Osnabrueck: 21-24 October 2010

Reuven Tsur

The Poetic Mode of Speech Perception Revisited

What our Ear Tells our Mind

Stating the Problem

This paper is an attempt to integrate (with some innovations) what I have said during the years about the rich precategorial auditory information reverberating in the background while we read poetry. Though in my age everything I say relies on my work during the past decades, I feel I have a few new insights too. And, I also feel it's time to bring together all that stuff into one corpus.

My work on this topic draws upon two different sources. One source is Anton Ehrenzweig's seminal work on Gestalt-free and thing-free qualities in the visual arts and music, and the interaction of such qualities within and across the boundaries of gestalts. When boundaries are clear-cut, colour interaction is increased *within* them and inhibited *across* them; the more blurred the boundaries, the stronger the interaction *across* them. Ehrenzweig discusses this via colour induction in the visual domain, and overtone fusion in polyphonic music. Overtone fusion in music may generate hitherto unheard tone colours, and enhance the gestalt-free *texture*. In speech, vowels are uniquely determined by concentrations of overtones called "formants". I claim that in certain circumstances the musical effects of poetry are crucially affected by similar overtone fusion. My other source is, obviously, speech research, which explores the transmission of speech through a stream of precategorial sound information, subsequently recoded into a sequence of speech categories.

Traditional literary scholarship has explored the versification devices which render poetry more musical than prose: metre, rhyme, alliteration, etc. In this paper I propose to go two steps beyond that. What I propose to explore is quite elusive, and traditional scholarship doesn't even have a vocabulary to refer to it. And even when I propose one, it will be impossible to define the conditions in which the terms apply. But, I hope, they will enable us to discuss elusive intuitions in a meaningful way.

As a first approximation, let us make the following distinction: Sometimes we experience the sound patterns of poetry as relatively opaque speech categories; and sometimes as abounding in rich precategorial auditory information reverberating in the back of our mind—in other words, alliterations may "click" or "clink". The reverberating background texture sometimes acts in a way that is similar to the gestalt-free shadings, scribblings and slight variations of color in the background of visual designs. They foreground the speech categories and sound patterns, round them out, as it were, making them more plastic and plump. Let me clarify what I mean by "resonance", "lingering auditory information" and "reverberating overtones", through an example adapted from Leonard Bernstein: "Depress middle C very carefully so as not to let it sound; then sharply strike and quickly release the

C an octave below. As soon as the lower C is released what will you hear? The *upper* C! It seems like magic, because you have really not "struck" this higher C, but the lower one" (Bernstein, 2004: 198), exciting the upper C-string to vibrate sympathetically as the first overtone of the C an octave below [`listen`]. Such activation of overtones is called resonance.

As a kind of "ostensive definition" of what I mean with reference to verbal structures, let me give three brief examples. First, a most elementary, nonpoetic example. Consider the name of the German philosopher *Kant*, and the English word *can't* (contraction of *cannot*). In the former, the [n] is a full consonant; in the latter it is attenuated into the [+NASAL] feature of the nasal vowel [ã], In the former it is perceived as relatively opaque; in the latter as more resonant than either a nasal consonant, or an oral vowel (e.g., [a]). Second, consider the following stanza from FitzGerald's The Rubáiyát of Omar Khayyám:

> Some for the Glories of this World; and some
> Sigh for the Prophet's Paradise to come;
> Ah, take the Cash and let the Credit go,
> Nor heed the rumble of a distant Drum!

Consider the three rhyme words, *some–come–Drum.* Some readers report that they are aware of a rich body of reverberating auditory information in *Drum*, but are not aware of a similar richness in the preceding rhyme-fellows. To be sure, traditional criticism has an excellent explanation for this, as far as it goes. There is an exceptionally rich alliteration pattern in ***rum**ble–**d**istant–**Drum***. The phrase refers to reverberating sound; the consonants [r] and [m], in turn, are perceived as *somehow* imitating sounds in nature. The present paper purports to go two steps further, and invoke the rich precategorial auditory information on the one hand, and the fusion of such auditory information on the other. The point is that in certain circumstances such resonance is enhanced, and in some inhibited. In this stanza of the Rubáiyát there are additional alliterations, though less resonating: ***Pr**ophet's **Par**adise* and ***C**ash–**C**redit*. Intuitively, they are perceived as less reverberating, more "opaque", and having a "leaner body".

Third, let us consider another classical example, Tennyson's notorious verse line "And murmuring of innumerable bees". It contains the sound cluster *mər* three times: twice in *murmuring,* and once in *innumerable.* Now, consider John Crowe Ransom's transcription of this line: "And murdering of innumerable beeves"—the reverberating background texture disappears. Ransom's transcription contains the sound cluster *mər* only twice; the rich precategorial information associated with it still could reverberate in acoustic memory (just as in *can't*), but it doesn't. Obviously, the onomatopoeia disappeared too. I will return to these examples. At the present stage of my argument I only want to point out one more thing. The [b] of *innumerable* and *bees* (or *beeves*) is part of another alliteration pattern, but not of the

onomatopoeia. But even the [b] seems to have a fuller, richer, more resonant body in Tennyson's line than in Ransom's rewriting. In course of this paper we shall encounter a wide range of conditions that may enhance or inhibit the reverberating sound information.

Such "ostensive definition" assumes an intuitive understanding, that is to say, that participants have sufficiently grasped the phenomenon to recognize the type of information being given. Alternatively, the vocabulary and theoretical framework to be expounded here may be useful in directing attention to certain elusive aspects of the sound dimension of a poem. Paraphrasing Morris Weitz, it can be used as a crucial recommendation what to look for and how to look at it in a given sound pattern.

The catch is that if someone cannot hear what I attribute to those examples, I cannot argue with him, nor bring him any proofs. Some people may raise an eyebrow at such an ostensive definition of my elusive topic. According to Frank Sibley (1962: 77), however, that is precisely how aesthetic concepts are and should be handled. "If we are not following rules and there are no conditions to appeal to" Sibley says, "how are we to know when they are applicable? One very natural way to counter this question is to point out that some other sorts of concepts also are not condition-governed. We do not apply simple color-words by following rules or in accordance with principles. We see that the book is red by looking, just as we tell that the tea is sweet by tasting it. So too, it might be said, we just see (or fail to see) that things are delicate, balanced, and the like"; or, "reverberating" at that. Or, as Manfred Bierwisch (1970: 108) says, poetics must accept effects as given.

Another vantage point to approach this phenomenon is from the Jakobsonian model of language functions. From the reader's point of view, there is a hierarchy of *arbitrary* signs: graphemes→ phonemes→ meaning→ extralinguistic referent (each later item being the signified of the preceding one). Man, as a sign-using animal, is programmed to reach the extralinguistic referent as fast as possible. According to Jakobson, what differentiates the referential function is focusing on the extralinguistic context; the poetic function focuses on the message. Figurative language directs attention to the semantic component of language, whereas the patterning of speech categories (versification) to the phonetic component. I have elsewhere discussed picture poetry, that forces the reader to attend back to the patterning of graphemes —hence perceived as so "artificial".

From the listener's point of view, speech sounds are transmitted by a stream of rich precategorial auditory information, which is immediately recoded into phonetic categories, and excluded from awareness. We only perceive a unitary, discrete phonetic category as [i] or [u]. Some of the precategorial auditory information, however, lingers on subliminally in active memory, and is available for certain cognitive tasks and aesthetic effects. Such lingering auditory information normally serves to preserve verbal material in active memory for efficient processing. It is active, usually, in the background, unnoticed. The present suggestion is twofold: first, that in poetic language, some or many listeners attend back not only to the

patterning of speech categoies explored by traditional rhetoric and criticism, but also to the lingering precategorial auditory information, turning it to aesthetic end in that it is perceived as musicality, onomatopoeia, or expressive sound patterns; and second, that in some circumstances more of the reverberating precategorial auditory information is perceived, and in some—less.

Some Experimental Evidence

There is plenty of experimental evidence for the reverberation and interaction of lingering auditory information. I will briefly mention only two sets of experiments. Let us turn to a set of experiments conducted for a different purpose. Researchers at the Haskins Laboratories (e. g., Liberman and Mann, 1981: 128-129; Brady et. al., 1983: 349-355; Mann, l984: 1-10), investigated the possible causes of some children's difficulty to learn to read, and revealed a deficiency in the use of phonetic coding by poor readers; good readers, by contrast, seem to make an excellent use of it. In one experimental task, poor readers had greater difficulty than good readers in tapping once or three times in response to the number of syllables in such spoken words as *pig* or *elephant,* or once, twice or three times in response to the number of phonemes in such words as *eye, pie* or *spy.* This has been interpreted as a deficiency in the use of phonetic coding. In another task, they had to memorize groups of words—either rhymed or unrhymed, as in the following ones:

chain	train	brain	rain	pain
cat	fly	score	meat	scale

Good readers did consistently better with both kinds of groups than poor readers. However, with the rhymed groups, their performance seriously deteriorated. While their reliance on phonetic representation increased their overall performance, the similar sounds of the rhyming words reverberating in their acoustic memory seem to have caused confusion. Good readers made efficient use of phonetic coding, whereas the poor readers made inefficient use of the acoustic information in short-term memory, and so were not penalized by the similar sounds of the rhyming words.

The sound patterns of poetry in general, and rhyme in particular, typically exploit the precategorial acoustic information and, actually, enhance its memory traces. In nonaesthetic memory experiments, this reliance on phonetic representation reveals two typical effects. It enables verbal material to linger for some time in short-term memory for more efficient processing, but also may cause acoustic confusion in certain circumstances.

The disadvantage of efficient readers with rhymed words seems to contradict our commonsense observation that versification facilitates the memorization of texts. But the contradiction is only apparent. In the experiment, the effect depends on the distance between the rhyming words. As Crowder (1983: 255) suggested in the set of experiments quoted below, "if the two units are too close together, they will integrate rather than inhibit. If they integrate, the subject will lose valuable informa-

tion". In poetry, by contrast, the rhyme words are further away from one another; break up a longer text into easily-remembered chunks; at the same time, the reverberating similar sounds unify the segmented text, and also enlist auditory memory in the service of remembering. When in the experiment the rhyme words come in close succession in a meaningless list, there is no intervening text to organize, nor is there a meaningful context that would impose semantic or grammatical constraints: it makes no difference which word comes first, which comes next; so, the fusion of formants becomes mere confusion. What in the nonaesthetic memory experiments is called acoustic *con**fusion**,* in an aesthetic context co-occurs with a coherent text, and may be perceived in the background as "harmonious ***fusion***", "musicality".

Experimental literature suggests three possibilities in the perception of successive speech stimuli. If a subsequent stimulus is very similar to the preceding stimulus, it may generate an enhanced response, because of integration with the lingering auditory information; if it is moderately similar, it will be reduced, inhibiting the lingering auditory information ("lateral inhibition"); if there is no similarity, it will be unaffected. In ordinary verbal communication usually one of the latter two possibilities is the case (Crowder, 1982, a-b). As to rhyme or FitzGerald's "rumble–Drum" alliteration, obviously the "very similar" option is the case. Robert G. Crowder suggests (personal communication) that there would be precedent for the assumption that the total effect would be the larger for having had a repeated sound. This depends on his assumption that both inhibitory and enhancing interaction takes place within the formant energy of the words, even though they may be spoken at different pitches (formants are concentrations of overtones that uniquely determine vowels). Thus, such sound patterns as rhyme and alliteration not only "exploit" the working of the auditory short-term memory, but actually enhance it.

A chapter in one of my books is called "Musicality in Verse, and Phonological Universals". I took the first part of this phrase from Kenneth Burke, the second from Roman Jakobson, and combined them. It took months before I discovered that I had created a most powerful alliteration: the word "verse" recurs entirely in the word "uni*vers*als". It would appear that in prose discourse our pronunciation of the same sound sequence in two words tends to be moderately similar, so as to reduce the lingering auditory information, directing attention away from the sounds to the referents. One may attend back to the alliteration by lengthening the sequence [vɜrs] of "uni*vers*als" and slightly raising its pitch, so as to render it more similar to "*vers*e". Likewise, the following sentence occurs in Ehrenzweig's account of Chevreuil's colour-induction experiment quoted below: "On a *gr*een *gr*ound the *gr*ey square would turn a distinct pink". I have quoted the passage numberless times during the past four decades or so, but only now, when writing the present paper I noticed the exceptionally powerful alliteration.

Crowder raises the question what the lateral inhibitory process is good for. "In vision, a system of recurrent lateral inhibition […] has the obvious adaptive consequence of edge-sharpening. Something quite similar may go on in speech perception. For example, in rapid fluent speech, people rarely achieve the 'target

values' of vowels, in terms of formant frequency. A system that could enhance the discriminability of adjacent vowels with high spectral overlap would be handy, especially if it operated at a very early, sensory, level of processing" (Crowder, 1983: 256). For our purpose, probably another function of lateral inhibition in speech perception would be to prevent the sound stratum from distracting attention from the extralinguistic referent in the referential function. And conversely, we must carefully articulate the 'target values' of formant frequencies, if we want them to enhance rather than inhibit the lingering auditory information. This may illuminate the phonetic mechanism underlying the "poetic function".

Speech Mode, Nonspeech Mode and Poetic Mode

Speech researchers distinguish between a speech mode and a nonspeech mode of auditory perception, which follow different paths in the neural system. In the speech mode there is typically a lack of correspondence between acoustic cue and perceived phoneme: we hear a unitary phoneme that is very different from the stream of auditory information that conveys it. In the nonspeech mode (natural noises, music, sonar etc.), by contrast, the shape of the perceived sound is similar to the shape of the sound wave. We seem to be tuned, normally, to the nonspeech mode; but as soon as the incoming stream of sounds gives the lightest indication that it may be carrying linguistic information, we automatically switch to the speech mode. In certain artificial laboratory conditions we may hear the phoneme in one ear, and the inarticulate noise in the other. We may also *see* it by converting speech into colour patches in images called sonograms or spectrograms.

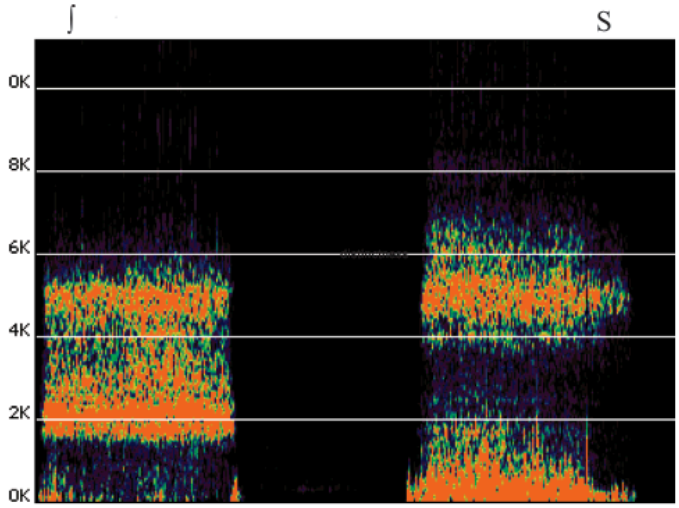

Figure 1 Sonograms of [ʃ] and [s], representing the first and second formant, and indicating why [s] is somehow "higher"

I have suggested that there is a third, "poetic mode". In the poetic mode, you hear the phonetic categories as in the speech mode; but some of the lingering precategorial auditory information becomes available too. Pronounce [ʃ] and [s] and try to determine which one is higher. Most people find that [s] is higher. Or pronounce [i] and [u] on the same pitch; you will probably find that [u] is lower and darker than [i]. As figures 1 and 2 show, the second formant of [s] and [i] is higher than that of [ʃ] and [u]. The first two formants of [u] are nearer together than those of [i] and, as Delattre et al. (1952) demonstrated, the human ear can even be fooled into hearing an [u] carried by two formants, when, in fact, generating one formant at an intermediate pitch. This indistinctness of [u] is perceived as darkness, as it were.

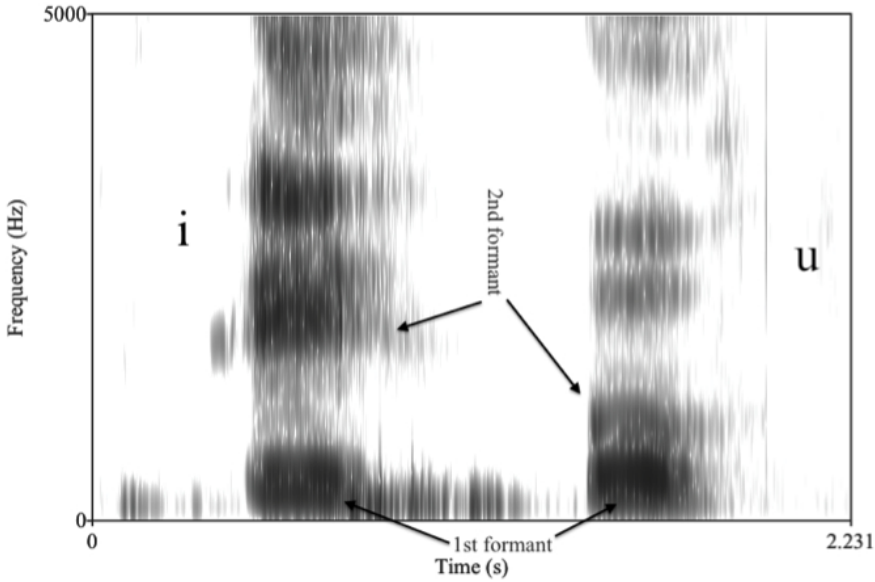

Figure 2 Spectrograms of [i] and [u]. The second formant is higher and the distance between the first two formants greater in [i] than in [u].

We must consider yet another distinction, that of relative "encodedness". Ask someone to pronounce [ba], [da] and [ga] on the same pitch and tell which one is higher. Not as with [s] and [ʃ] or [i] and [u], most (but not all) people will have difficulty to tell this. The only difference between [ba], [da], and [ga] is the onset frequency of the so called second formant transition, in this ascending order. But stop consonants are highly encoded, that is, little or no lingering auditory information reaches awareness.

One may make two successive distinctions in the acoustic structure of speech sounds. Some speech sounds (as [p, t, k]) are abrupt; some are continuous. Continuous speech sounds may be periodic (as [l, m, n, j]; or aperiodic (as [ʃ, s, f]). [r] is continuous, periodic, and multiply interrupted. Abrupt sounds are usually highly

encoded; continuous aperiodic sounds somewhat less encoded, and periodic sounds relatively unencoded. Even among vowels, as we have seen, nasal vowels are less encoded than oral vowels. Voiced stops are an interesting case in point: stops are abrupt, whereas voicing is periodic. I have used above the pair of onomatopoetic verbs *click* and *clink*. The word-final [k] in *click* is exceptionally sharp and abrupt; the nasal vowel in *clink* is periodic and reverberating, and quite readily available to awareness.

Coarticulation increases the encodedness of speech sounds. There is experimental evidence that in isolated vowels and continuants more precategorial information reaches awareness than when pronounced with another speech sound (Rakerd, 1984; Repp, 1984). In [ʃ. s] it is easier to discern which one is higher than in [ʃa, sa]. Symbolist poets sometimes capitalize on this fact, enhancing the reverberation of speech sounds. In Rimbaud's "Voyelles" the vowels [a. e. i, ü, o] are directly named, with no consonantal context, yielding a stream of acoustic energy. The Hungarian poet Kosztolányi wrote a poem inspired by his wife's name, *Ilona,* all continuous periodic sounds. In one of the stanzas he enumerates the isolated sounds of her name: "Csupa *l*,/ csupa *i*,/ csupa *o*, csupa *a*"—all continuous streams of periodic sounds.

We have been exploring the question, why in some poetic contexts one may discern reverberating auditory information in the background, bestowing on the speech sounds plasticity and a "fuller body", whereas in other contexts speech sounds are perceived as relatively "lean" and sharply defined. Lingering precategorial auditory information is the clue. One distinction we have made concerns "encodedness". In some speech sounds the precategorial auditory information is more readily available than in others.

This by itself, however, would be a quite rigid phenomenon, insufficient to account for the experience we set out to explain. But, in certain circumstances, readers and poets may switch attention from one aspect of the speech sounds to another. Consider our example "And murmuring of innumerable bees" as opposed to "And murdering of innumerable beeves". I have suggested that the [b] of *innumerable* is no part of the onomatopoeia, but even the [b] seems to have a fuller, richer, more resonant body in Tennyson's line than in Ransom's rewriting. In *murmuring* and *innumerable* the meaning directs attention to the rich precategorial auditory information available in the continuous periodic consonants [m], [n], [r], and [l]. The voiced stop [b] is perceived in the *murdering* context as a unitary speech sound that typically blocks, so to speak, the passage of reverberating acoustic energy. The *murmuring* context, by contrast, separates the periodic "voiced" feature in [b] and activates it, blending its voiced, periodic element with that of [l]. Such an "aspect-switching" will appear less incredible if we note that one of the most effective cues for voiced consonants is an articulatory gesture plus voice onset time: that is, how much time passes between the articulatory gesture and the beginning of voicing.

There are indications that this is not a mere freak of an artificial rewrite exercise. Thus, for instance, Iván Fónagy (1961), in his study of the expressiveness of speech sounds compared the relative frequency of phonemes in ten especially angry poems and ten especially tender poems by a variety of poets in French, German and Hungarian. Such voiceless stops as [p, t, k] are "angry" for most poets. I interpreted this fact as having to do with their abrupt and highly encoded nature. Tender moods are more open to uncategorized sensory information than aggressive moods. Now consider the relative frequency of /g/ and [d] in Victor Hugo's and Paul Verlaine's poems. /g/ occurs over one and a half times more frequently in Verlaine's tender poems than in his angry ones (1.63: 1.07), whereas we find almost the reverse proportion in Hugo's poems: 0.96% in his tender poems, and 1.35% in his angry ones. As to /d/, again, the same sound has opposite emotional tendencies for the two poets, but with reverse effects. For Verlaine it has a basically aggressive quality (10.11: 7.93), whereas for Hugo it has a basically tender quality (7.09: 5.76)—again, in almost the same reverse proportion. The reason seems to be similar to the one we have offered for the shift of perceived qualities of [b] in the *murmuring* and *murdering* contexts: "aspect-switching". If you attend to the [g] or [d] as a unitary abrupt stop consonant, it may have a strong aggressive potential; if you attend to the periodic voiced ingredient, it may contribute to a tender quality. Obviously, Hugo and Verlaine applied to these voiced stops the same cognitive mechanism, but with a reverse focus.

Colour and Overtone Interaction

As I mentioned above, in this paper I make an attempt to apply to poetry Anton Ehrenzweig's work on colour and overtone interaction in the visual arts and music. According to the gestalt psychologists, "colour interaction increases within the boundaries of a good gestalt while it is inhibited across its borders" (Ehrenzweig, 1970: 172). In the visual domain, the process was demonstrated experimentally, in a most dramatic way, back in the early nineteenth century. Ehrenzweig (1970: 170–171) describes a demonstration of colour induction by Chevreuil (which I attempted to replicate in figures 3, 4 and 5). "The experiment which demonstrated interaction most clearly was to place a small grey square on a large ground of colour. On a green ground the grey square would turn a distinct pink". I have found that it takes the square a few seconds to turn pink. "A few years later a most paradoxical phenomenon was observed; when a sheet of semi-transparent tissue paper was placed over the whole area the saturation of the green ground was of course severely diminished. One would have expected that the colour induction in the grey would be reduced to the same extent, that is to say that the induced pink of the grey square would also become much paler. But the opposite happened: the pinkness of the grey square became more pronounced". It was Helmholz who found an explanation for this paradox: the tissue paper made the outline of the grey square fuzzier and this weakening of its form increased colour interaction across its boundaries. "A

comparatively crude weakening of the line was sufficient to compensate—indeed more than compensate—for the enormous loss in the saturation of the colours," says Ehrenzweig. "As in all relationships between form and colour the reverse effect can also happen. Strong colour interaction tends to make sharp outlines seem much softer than they are".

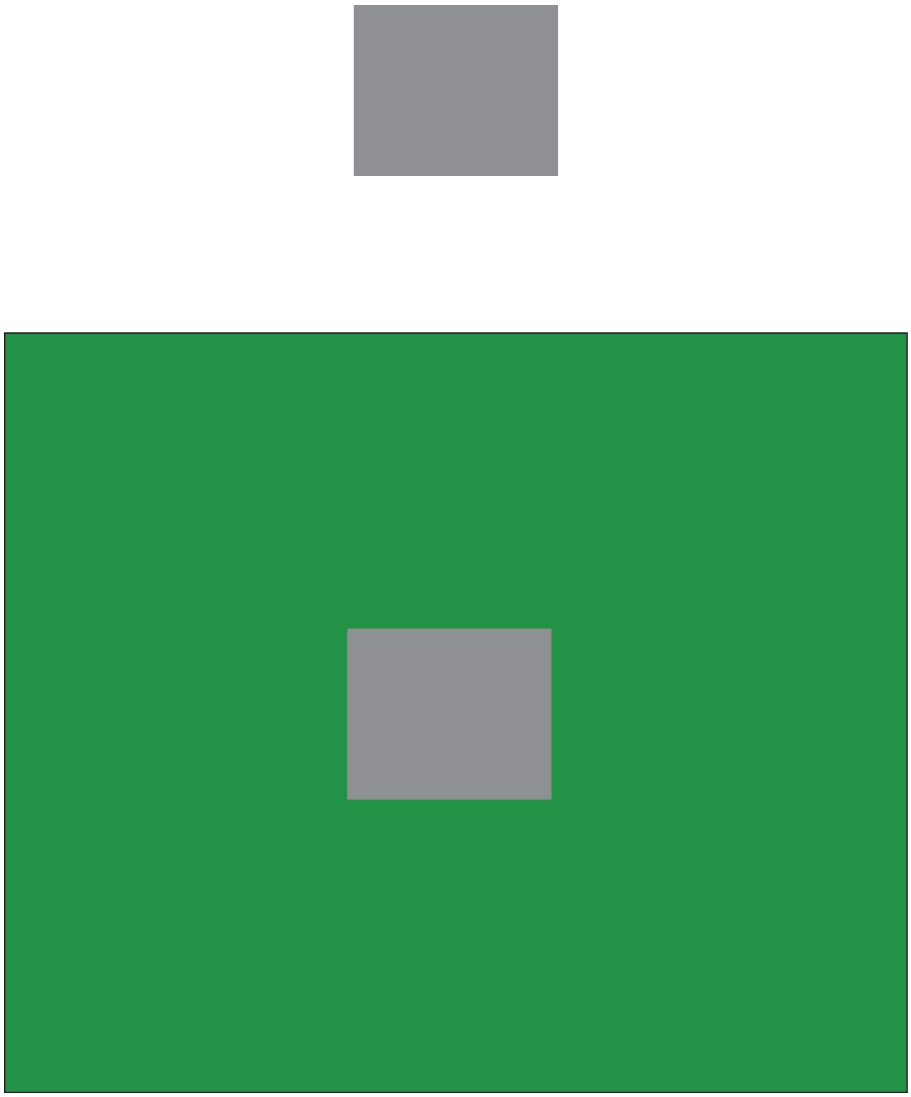

Figure 3 Chevreuil's demonstration of colour induction replicated: A gray square on a green ground gradually turns pink. The small square on the green ground is exactly the same as the one on the top of the figure.

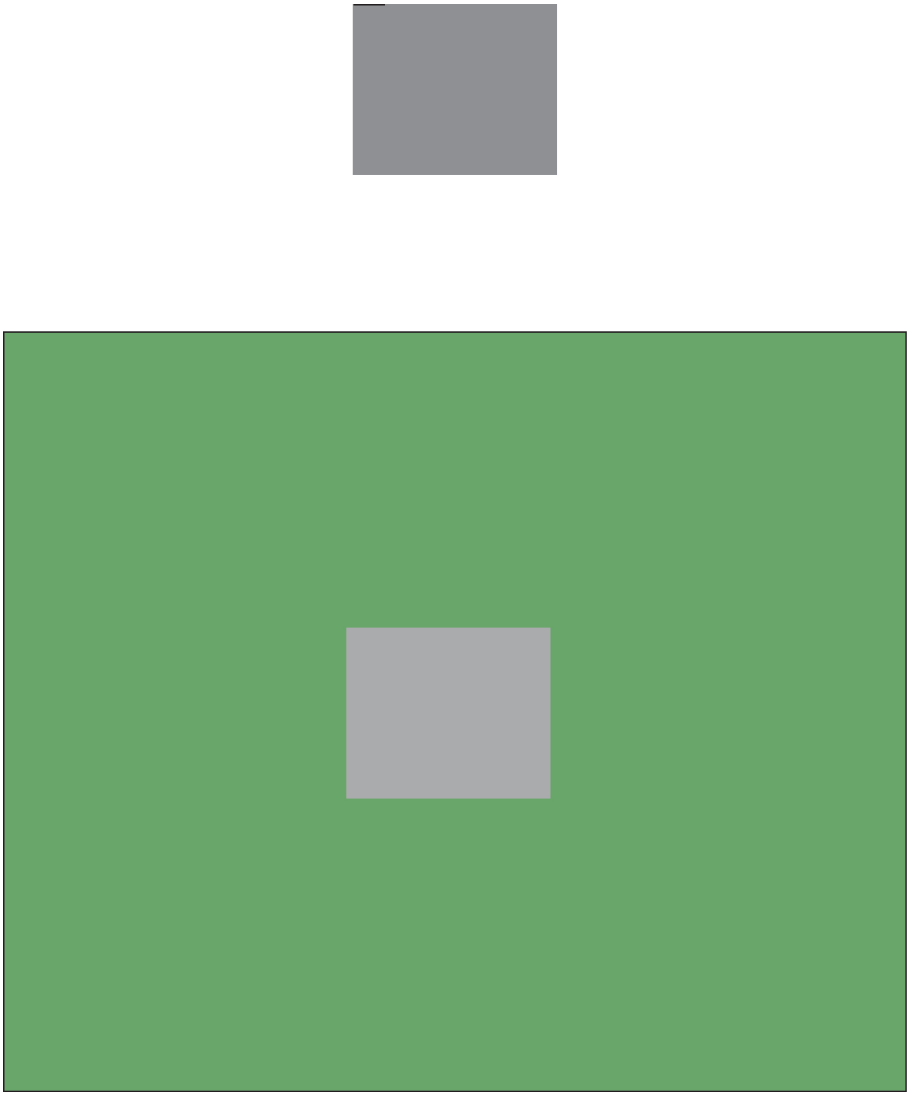

Figure 4 When covered with a semitransparent tissue paper, colour induction is increased rather than decreased. Here the tissue-paper effect was simulated by the Opacity function of Adobe Illustrator.

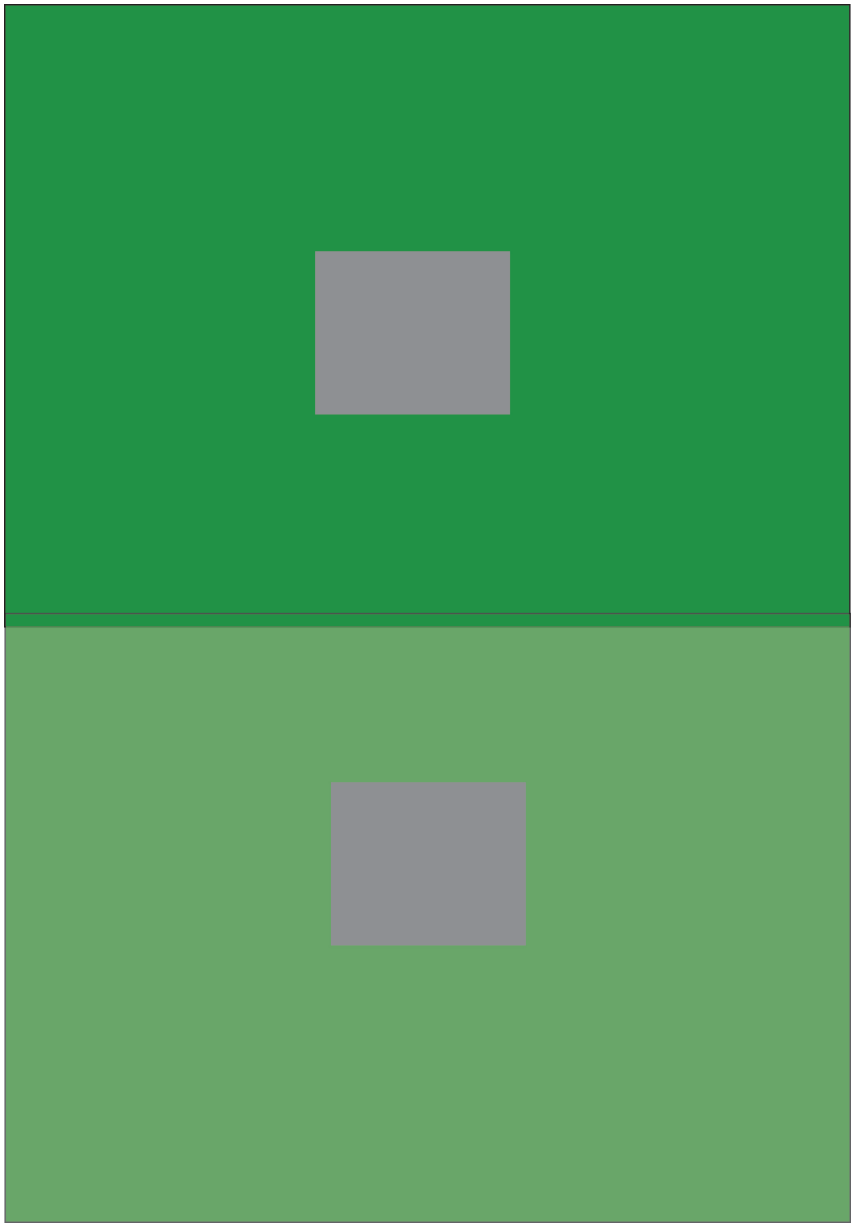

Figure 5 The more and less opaque versions side by side. In the lower field, the pink of the small square is more pronounced.

Ehrenzweig argues that the same relationships hold in music between melodic shapes and overtone interaction (overtones being the physical correlates of "sound colour") (Ehrenzweig, 1965: 172–173). This is more immediately relevant than colour induction to poetry reading, where we are dealing with intonation shapes and overtone interaction (where overtones are the physical correlates of vowels). "To the extent to which a musical note is fitted into a clean melodic 'line' it is prevented from fusing into harmonic tone 'colour'; conversely a strong chord will temporarily fuse the loose strands of polyphony into solid tone colour so that the separate melodic lines disappear altogether. I have mentioned that the ear constantly oscillates between the harmonic fusion and polyphonic separation of the melodic lines; this conflict between 'form' and 'colour' belongs to the very life of music" (1970: 173). There are good reasons to suppose that a similar dynamics, with the necessary changes, may occur in poetic language, in respect of the lingering rich precategorial auditory information on the one hand, and, on the other, clear-cut speech categories and good syntactic and prosodic gestalts.

Thus, one of the several ways to manipulate the precategorial auditory information in poetry concerns gestalts and boundaries. Stronger boundaries would inhibit interaction across them, the weakening of boundaries may boost it. That may be one difference between the perceived quality of Pope's and Milton's alliterations.

At this point we may try to account for some of the sound effects in FitzGerald's stanza quoted above. Consider the phrase "***rum**ble* of a ***d**istant **Drum***". The repeated sound cluster [rʌm] + [l} are lowly-encoded speech sounds, that is, much of the rich precategorial auditory information may reach awareness. In light of Crowder's experiments with lateral inhibition, in such repeated sound clusters, the similar formants may be integrated in certain circumstances, and enhance each other. Furthermore, ***rum**ble* is defined by the Merriam-Webster Collegiate dictionary as "a low heavy continuous reverberating […] sound", and may direct attention to the reverberating periodic background texture of these phonemes. According to the foregoing discussion, it may also direct attention to the periodic ingredient in the voiced stops [b] and repeated [d]. That is why this phrase tends to be perceived as reverberating, as having a "full body".

Or, consider the [r] in ***r**umble, d**r**um, P**r**ophet, Pa**r**adise* and *C**r**edit*. The [r], as I said, is periodic and multiply interrupted, and lowly-encoded. Thus, the reader/listener has the option to switch between three aspects of the sound, according to the context: the unitary speech sound; its interrupted (that is, abrupt) nature, sometimes suggesting threat, aggression or frightening noises; and its periodic nature, suggesting a soft resonant quality (as in *murmuring*). In *P**r**ophet, Pa**r**adise* and *C**r**edit* [r] is associated with the unvoiced stops ([p] and [k]), which are likely to direct attention away from its resonance to the unitary speech sound.

Elsewhere I argued that reverberating precategorial auditory information is inhibited if the text is under control of a strong gestalt. If the gestalt is weakened, reverberation increases; if it is improved, reverberation decreases. Consider the alliteration in the line "Ah, take the **C**ash and let the **C**redit go". Here the unvoiced [k] sound is repeated, which is highly encoded, that is, little or no precategorial

information reaches awareness. The alliterating syllables ***C**ash* and ***C**redit* are stressed, foreground the strong positions and enforce the symmetrical structure, strengthening the overall gestalt. Even if *some* precategorial sound information could be perceived, reverberation would be contained. Indeed, the alliteration sounds relatively dull, hard and compact. Compare now this alliteration with another repeated [k], in No 3 of the *Rubáiyát*:

And, as the Cock crew, those who stood before

Here two consecutive stressed syllables blur the metric pattern. We are on slippery ground. But it would appear that the alliteration in *Cock crew* sounds less compact than in *Cash* and *Credit*, To be sure, the word *crew* (but not *Cash* and *Credit*) refers to a loud shrill sound, and thus could activate whatever inarticulate sound information is available in the highly-encoded [k] and the lowly-encoded, multiply interrupted [r]. But notice this. Manipulating the second alliterating stressed syllable into a strong position strongly enhances metre and the alliteration may be perceived as somewhat more tightly packed, less resonant.

And, as the Cock that crew — who stood before

Now suppose all these differences reflect a mere whimsey, my idiosyncratic impression, and doesn't conform with your intuition. The present apparatus enables me at least to communicate what I mean; and allows you to know what it is you are expected to but don't hear. We *hear* (or fail to hear) that *Cock that crew* is less resonant than *Cock crew*. But the present apparatus renders the topic *discussible*, And in some instances, at least, my argument may even help to direct your attention to what I hear.

Or take the phrase *Prophet's Paradise*. Among the unvoiced stops, [p] is phonetically softer, less compact than [k]. But the proximity and syntactic connection between the two words enhances their popping noise, and the intervening unstressed syllable in a weak position renders the *pr* alliteration relatively nonresonant. We have been discussing the version of the third and fourth edition of the *Rubáiyát*. In the second edition FitzGerald had:

Some for the Glories of This World; and some
Sigh for the Prophet's Paradise to come;
Ah, take the Cash, and let the Promise go,
Nor heed the music of a distant Drum!

By substituting *Credit* for *Promise*, FitzGerald improved by one masterstroke both the parallelism of figurative language and the alliteration pattern. In the earlier version, ***P**romise* joins the ***P**rophet's **P**aradise* alliteration pattern as a third member. However, it would appear that ***P**romise* is less obtrusive than the other two

members; the intervening line and clause boundary seems to inhibit acoustic interaction across it. “Music of a distant Drum” too directs attention to the auditory element in meaning. But notice this: *music* shares with *rumble* the stress pattern, and the meaning component [+sound]. But the focus of the alliteration shifts from *rumble–Drum*, to ***d**istant **D**rum,* foregrounding the abrupt [d] rather than the reverberating [rʌm]. Consequently, the repeated [d] changes its character: one tends to attend to the harder unitary phoneme rather than to the reverberating, periodic voicing. Thus, it is perceived more as a sound imitation of “beating drumsticks” than of "reverberating membrane".

Alliterations like “Ah, take the **C**ash and let the **C**redit go” or *Prophet's Paradise* would be typical of Pope, while alliterations like *Cock crew* would be typical of Milton, and quite untypical of Fitzgerald or Pope.

This may also explain why Milton and Shelley, two of the most deviant poets in English, are thought to be two of the most musical poets too. Their frequent strings of stressed syllables and other deviations blur the metric gestalt; their frequent enjambments blur the gestalt of the verse lines. This may background the sound patterns, but increase their overtone interaction over considerable stretches of text.

According to Jakobson, rhyme and alliteration focus attention on the sound stratum of a text; the present conception further distinguishes between two layers within the sound stratum: the string of phonetic categories and stream of precategorial auditory information. Traditional criticism typically deals with the string of phonetic categories. The present work assumes that by moving from the phonetic categories to the precategorial sound information one does not merely restate the issues once in “phonetics language” and once in “acoustics language”, but crucial distinctions can be made on the precategorial level that are not automatically implied by the phonetic categories. It propounds a vocabulary and theoretical framework that enable us to handle the precategorial auditory information in a principled manner. In the present case, for instance, in onomatopoeia it may point out and account for the more fine-grained interaction of sound and meaning. Moreover, it may offer *several* interacting reasons why in “rumble of a distant Drum” the precategorial auditory information is more obtrusive, more resonating, than in its rhyme fellows or some other alliteration patterns in the stanza.

Individual differences

The experiments with efficient and poor readers may suggest that there are individual differences in the handling of precategorial auditory information. I will point out two kinds of reasons for such individual differences: the ones related to phonetic coding, and those related to personality style. From the afore-mentioned experiments we know that some persons have a deficit in phonetic coding; but it doesn’t say what coding they are using instead. Crowder and Wagner (1992: 228-230) summarize an experiment by Byrne and Shea which strongly suggests that they are using a semantic code. In this experiment, subjects had to take a “reading test”,

reading out lists of words and then, unexpectedly, were given a memory test. They were presented with the words read earlier, interspersed with a number of additional words, to which they had to respond "old" or "new". The "new" words were either phonetically or semantically related to the "old" words. "Assume the prior items were *home* and *carpet: house* and *rug* would be the semantically similar foils and *comb* and *market* would be the phonetically similar foils". Good readers tended to confuse both phonetically and semantically related words, poor readers semantically related words. This would suggest that good readers use both phonetic and semantic coding, poor readers mainly semantic coding.

There are personality styles that are intolerant of unique, unclassifiable sensations, which too may penalize some persons in responding to precategorial auditory information. Consider the following description: "The leveler is more anxious to categorize sensations and less willing to give up a category once he has established it. Red is red, and there's an end on't. He levels (suppresses) differences and emphasizes similarities in the interest of perceptual stability. For him the unique, unclassifiable sensation is particularly offensive, while the sharpener at least tolerates such anomalies, and may actually seek out ambiguity and variability of classification" (Ohmann, 1970: 231).

The foregoing line of investigation may have tapped a kind of individual differences in verbal strategies that may characterize very advanced readers too; I mean students and professors of literature. Such measures as the ones applied to first graders would be insufficient to measure a literature student's relative reliance on phonetic coding. However, we may find in prosody classes students of literature who are incapable of telling which syllable of a bisyllabic word in their mother tongue is stressed, even though they can pronounce it correctly. This suggests that they can spontaneously retrieve the word from long-term memory together with its sound pattern and stress pattern, as a compact package. But it does not reverberate in their processing space to enable them to inspect it from a higher standpoint. Some professors, in their published work, consistently prefer to rely on the meanings rather than the sounds of poetry. The plain fact seems to be that even at this level of students and professors of literature there are individual differences in this respect, and there are persons who do not seem to feel at ease with phonetic representation, and seek to fall back as frequently as possible on semantic coding. This is most conspicuous in the response of various readers to what Snyder (1930) called "hypnotic poetry". In a number of poems by Poe, Coleridge, and certain other poets, many readers are inclined to "attend away" from the meaning of the words, and to become "spell-bound" by their sound. The reader feels as if he were entangled by the sounds of these poems, and tends to perceive their meaning relatively dimly. However, some other readers respond quite differently: they may find the sound effects of the same poems rather boring and unemotional. Such differences may be due to a person's being high or low on the personality variable "absorption" (Glicksohn, Tsur & Goodblatt, 1991; Tsur, 2006: 130–137). Still other readers may ignore all in all the sound effects of the same poems and seek to account for their signifi-

cance solely in terms of their meaning. More often than not, such interpretations will also ignore the possible hypnotic effect of these poems. Such an inclination to account for the significance of poems solely in terms of their meaning may also be an indication of a basic preference for semantic rather than phonetic coding.

I have an illuminating piece of direct evidence for this tendency, on the professor level. In my research on the rhythmical performance of poetry I explored two recordings by John Gielgud of the last line of Shakespeare's Sonnet 129. I made Gielgud's readings available on my webpage and sent a request to the PSYART and Coglit lists to respond to them. I received all in all five fairly detailed responses. An outstanding psychoanalytic critic who had extensively written on this poem commented at the end of his response: "My attention was called to the rhythm of the second reading and it was somehow distracting and thereby detracting".

Hopkins's "The Windhover"

Gerald Manley Hopkins' poetry is notorious for its acrobatic play on phonetic categories. In this section I will explore one of those virtuoso constructions and the interaction of the underlying precategorial auditory information—first speculatively, and then in an instrumental study of the means by which a vocal performer may boost or inhibit that interaction. In a forthcoming paper I explore the possible contribution of such verbal acrobatics to an ecstatic effect. It invokes Bergson's distinction between perceptions that are "clear, distinct, juxtaposed or juxtaposable one with another, and tend to group themselves into objects"; and "a continuous flux, […] a succession of states each of which announces that which follows and contains that which precedes it". The former underlies "ordinary consciousness", the latter—what for our purpose may be called "altered states of consciousness". Here I will reproduce only some of the phonetic, not the semantic part of that discussion, and will refer only to the first two lines.

> I caught this morning morning's minion, king-
> dom of daylight's dauphin, dapple-dawn-drawn Falcon, in his riding

(this morning I caught [sight of] the minion or servant of the morning, [who is] the dauphin, or crown-prince,
of the Kingdom of Daylight—a falcon spotted or dappled by the dawn as he was riding—George P. Landow's "interlenear translation")

In view of Bergson's distinction between kinds of perception, both the meanings of words and the sound clusters denoting them are juxtaposed, and thus poorly suited to express an ecstatic experience. In view of our foregoing discussion, however, Hopkins' pileup of exceptionally similar sound clusters in close succession may have the following effect: it may reinforce the lingering rich precategorial auditory information in active memory, and may enhance its interaction, causing

fusion (or confusion), Thus, though the phonetic categories and their clusters are juxtaposed, "in reality no one begins or ends", in Bergson's terms, "but all extend into each other". Normally, as I said, rhymes break up a longer text into easily-remembered chunks. At the same time, the reverberating similar sounds unify the segmented text, and also enlist auditory memory in the service of remembering. In Hopkins' pileup of sound clusters, where there is no intervening text, the reverberating similar sounds extend into each other, fusing the discrete phonetic units, so as to contribute to the perception of a flux.

In what follows I will explore the vocal strategies by which a performer may boost or inhibit the interaction of such precategorial auditory information across word and phrase boundaries.[1] Richard Austin provides online a beautiful reading of Hopkins' "The Windhover" `[click to listen to it]`. I will play only the first two lines. In my forthcoming paper I deal with these two lines in great detail. Here I will focus on only one point. `[click to listen to the first two lines]`

I had a weird feeling that Austin's intonation leaps are too big at certain points, articulating too clearly certain word boundaries. This is by no means a mistake I am trying to correct. We seem to be at cross-purposes. While Austin's purpose seems to be to overarticulate boundaries so as to disambiguate the syntactic structure of obscure pileups, my purpose is to weaken boundaries, so as to boost overtone interaction. In my forthcoming paper I elaborate on several boundaries, but here I will focus on only one.

Figure 6 gives phonetic information about the phrase "dapple-dawn-drawn-Falcon", excised from this reading [`click to listen to it`]. Similarity of overtone concentrations (shown by spectrograms) obviously reflect the similar phonetic structures of words. Similarity of durations, pitch contours and amplitudes are largely due to the performer's manipulation.

The arrows in Figure 6 point at the first and second formant, at the point where the vowel and [n] meet in "dawn" and "drawn" (formants are concentrations of overtones that uniquely determine speech sounds; the first and second formants being the crucial ones). The shape of the three plots are very similar for the words "dawn" and "drawn", sufficiently similar to enhance the lingering precategorial auditory information in each other.

1 The sound files for this paper can be found online, at http://www.tau.ac.il/~tsurxx/Hopkins/tongue-twisters_sound_files.html

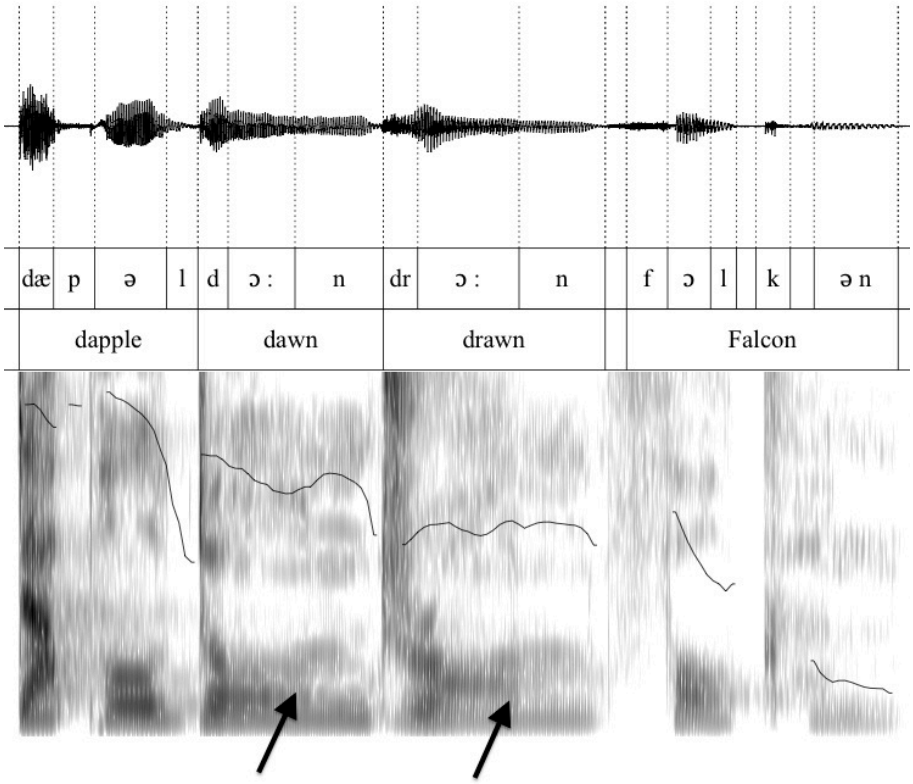

Figure 6 Wave plot, pitch contour and spectrogram of the phrase "dapple-dawn-drawn-Falcon", read by Richard Austin. The upper window presents the wave plot display, which shows a plot of the wave amplitude (in volts) on the vertical axis, as a function of time (in milliseconds) on the horizontal axis. The lowest window presents a spectrogram, which displays relative intensity as a function of time (horizontal axis) and frequency (vertical axis); as well as a fundamental frequency (pitch) plot, displaying time on the horizontal axis and the estimated glottal frequency in Hz on the vertical axis.

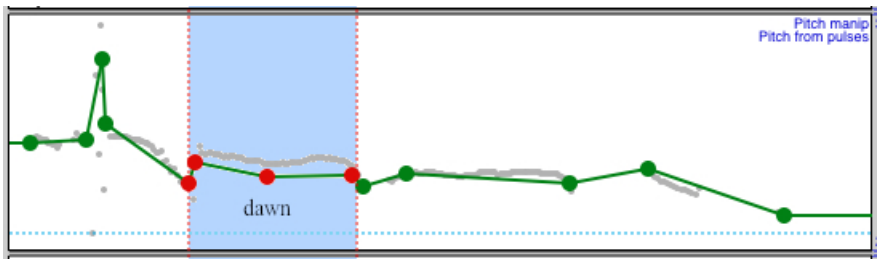

Figure 7 The genuine and "doctored" pitch contour of *dapple dawn drawn Falcon*. The higher, dim curve reflects the genuine reading, the lower curve with the "pitch dots" reflects the manipulated reading.

I had the impression that the distinct downstep from the pitch of "dawn" to that of "drawn" enhances the boundary between them, obstructing the interaction of lingering precategorial auditory information. So, I used the speech processor Praat to lower the pitch of "dawn" to equalize it with "drawn", blurring the distinct boundary between the two words, so as to boost the fusion of precategorial information [click to listen to it].The difference between the two versions will be better discerned when listening to them in close succession [`click to listen to it`]. Finally, let us listen to the first two lines of the poem with the doctored intonation contour substituted for the genuine one [`click to listen`]. In my forthcoming paper I manipulated two additional boundaries in these two lines. So, finally, let us listen to them with all the manipulated parts substituted for their genuine counterparts [`listen to it`].

Summary and Conclusions

In this paper I made an attempt to explore what our ear tells our mind when reading poetry. Speech researchers speak of a speech mode and a nonspeech mode of listening. I suggest that there is a third, poetic mode of speech perception too, where people attend to the clear-cut phonetic categories, but some of the inarticulate, precategorial auditory information does reach awareness. This auditory information has a mysterious effect on the perceptual and expressive qualities of the speech sounds. The business of this paper has been to explore an elusive aspect of this mysterious effect and the underlying cognitive mechanisms, epitomized as "some alliteration patterns we hear as 'click', some as 'clink'". The basic fact in such an endeavour is that there is a string of phonetic categories underlain by a stream of precategorial auditory information. In such a hierarchic structure there always lurks the danger of reductionism. The critic is prone to merely restate in acoustics-language (or brain-language) what could be said in phonetics-language. Cognitive Poetics is particularly susceptible to this danger. I claim that in this case there is no automatic one-to-one correspondence between the two levels of information. The precategorial information may affect, or fail to affect, awareness in a variety of ways. First, there are two context-independent ways. Some speech sounds are highly-encoded, some lowly-encoded; that is to say, in some speech sounds little or no precategorial information reaches awareness, in some—relatively much. Likewise, some personality styles are more, some less tolerant of unique uncategorized sensations. The most elementary context concerns co-articulation: Vowels in consonantal context are perceived as more encoded than isolated vowels; the same is true of continuants in vocalic context. As to the effect of the contents, the same lowly-encoded speech sounds may be heard with their precategorial auditory information reverberating in the background when referring to some sound or noise, or, in a neutral context may be perceived as a unitary phonetic category. There are techniques by which overtone interaction can be enhanced or inhibited, most notably by similarity, proximity, and gestalt boundaries. If a subsequent stimulus is

very similar to the preceding stimulus, it may generate an enhanced response; if it is moderately similar, lateral inhibition sets in. If the similar stimuli are too close together, in nonaesthetic memory experiments it may cause confusion; they will integrate rather than inhibit, and the subject will lose valuable information. Such confusion may become exceptionally effective musical fusion in an aesthetic context. Overtone interaction, just like colour interaction, increases within the boundaries of a good gestalt while it is inhibited across its borders. Both poets and vocal performers have their own techniques to weaken or strengthen the boundaries, so as to increase or decrease the interaction. If an alliteration pattern is under the control of a strong gestalt, it may keep down the resonance; the weaker the gestalts, the less restrained the reverberation of the precategorial sound information. Such voiced stops as [b, d, g] consist of an abrupt articulatory gesture and a stream of periodic voicing. Consequently, the reader or the poet has an option to switch between three aspects, as the context may demand: they may attend to a unitary voiced stop, or to the abrupt articulatory gesture, or to the periodic stream of voicing. [b] is perceived as more reverberating in a "murmuring" context than in a "murdering" context; Verlaine and Hugo attend to opposite aspects in [d] and [g], for opposite emotive effects. Consequently, speech sounds and repetitive patterns may be perceived as opaque and tightly packed together in certain circumstances; in other circumstances as plastic and freely reverberating. This paper has explored what reasons can be given to support one impression or another.

References

Bernstein, Leonard (2004) *The Joy of Music*. Pompton Plain N.J. & Cambridge UK: Amadeus Press.

Bierwisch, Manfred (1970) "Poetics and Linguistics", in Donald C. Freeman (ed.), Linguistics and Literary Style. New York: Holt, Rinehart & Winston. 97–115.

Brady, Susan, Donald Shankweiler & Virginia Mann (1983) "Speech Perception and Memory Coding in Relation to Reading Ability". *Journal of Experimental Child Psychology 35:* 345-367.

Crowder, Robert G. (1982 a) "A Common Basis For Auditory Sensory Storage in Perception and Immediate Memory." *Perception & Psychophysics 31:* 477-483.

Crowder, Robert G. (1982 b) "Disinhibition of Masking in Auditory Sensory Memory". *Memory & Cognition 10:* 424-433.

Crowder, Robert G. (1983) "The Purity of Auditory Memory". *Phil. Trans. R. Soc. Lond. B* 302: 251–265.

Crowder, Robert G, and Richard K. Wagner (1992) *The Psychology of Reading.* New York and Oxford: Oxford UP.

Delattre, Pierre, Alvin M. Liberman, F. S. Cooper & L. J. Gerstman (1952) "An Experimental Study of the Acoustic Determinants of Vowel Color". *Word 8:* 195-210.

Ehrenzweig, Anton (1965) *The Psychoanalysis of Artistic Vision and Hearing.* New York: Braziller.

Ehrenzweig, Anton (1970) *The Hidden Order of Art*. London: Paladin.

Fónagy, Iván (1961) "Communication in Poetry". *Word 17:* 194–218.

Glicksohn, J., R. Tsur and Ch. Goodblatt (1991) "Absorption and Trance-Inducing Poetry". *Empirical Studies of the Arts 9/2:* 115–122.

Landow, George P. (online) "A Reading of Gerard Manley Hopkins's 'The Windhover'". available: http://victorianweb.org/authors/hopkins/hopkins10.html

Liberman, Isabelle Y., and Virginia A. Mann (1981) "Should Reading Instruction and Remediation Vary with the Sex of the Child?" *Status Report on Speech Research SR-65:* Haskins Laboratories.

Mann, Virginia A. (l984) "Reading Skill and Language Skill". *Developmental Review 4:* 1-15.

Ohmann, Richard (1970b) "Modes of Order", in Freeman, Donald C. (ed.), *Linguistics and Literary Style*. New York: Holt, Rinehart & Winston. 209–242.

Rakerd, Brad (1984) "Vowels in Consonantal Context Are More Linguistically Perceived than Isolated Vowels: Evidence from an Individual Differences Scaling Study". *Perception & Psychophysics 35:* 123–136.

Repp, Bruno H. (l984) "Categorical Perception: Issues, Methods, Findings", in N. J. Lass (ed.), *Speech and Language: Advances in Basic Research and Practice, vol. 10*. New York: Academic Press. 243–335.

Sibley, Frank (1962) "Aesthetic Qualities", In Joseph Margolis (ed.), *Philosophy Looks at the Arts: Contemporary Readings in Aesthetics*. New York: Scribner. 63–88.

Snyder, E. D. (1930) *Hypnotic Poetry: A Study of Trance-Inducing Techniques in Certain Poems and its Literary Significance.* Philadelphia: University of Pennsylvania Press.

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

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

Tsur, Reuven (2008) *Toward a Theory of Cognitive Poetics*. Second, expanded and updated edition. Sussex Academic Press: Brighton and Portland.

Weitz, Morris (1962) "The Role of Theory in Aesthetics", in J. Margolis (ed.), *Philosophy Looks at the Arts*. New York: Scribner. 48–59.

Richard Austin Reads the Poetry of Gerard Manley Hopkins

http://victorianweb.org/authors/hopkins/windhover3.html