Studia Metrica et Poetica 8.1, 2021, 7–39

Acoustics and Resonance in Poetry: The Psychological Reality of Rhyme in Baudelaire's "Les Chats"

Reuven Tsur*

Abstract: This article uses the term "psychological reality" in this sense: the extent to which the constructs of linguistic theory can be taken to have a basis in the human mind, i.e., to somehow be reflected in human cognitive structures. This article explores the human cognitive structures in which the constructs of phonetic theory may be reflected. The last section is a critique of the psychological reality of sound patterns in Baudelaire's "Les Chats", as discussed in three earlier articles. In physical terms, it defines "resonant" as "tending to reinforce or prolong sounds, especially by synchronous vibration". In phonetic terms it defines "resonant" as "where intense pre-categorical auditory information lingers in short-term memory". The effect of rhyme in poetry is carried by similar overtones vibrating in the rhyme fellows, resonating like similar overtones on the piano. In either case, we do not compare overtones item by item, just hear their synchronous vibration. I contrast this conception to three approaches: one that points out similar sounds of "internal rhymes", irrespective of whether they may be contained within the span of short-term memory (i.e., whether they may have psychological relit); one that claims that syntactic complexity may cancel the psychological reality of "internal rhymes" (whereas I claim that it merely backgrounds rhyme); and one that found through an eye-tracking experiment that readers fixate longer on verse-final rhymes than on other words, assuming regressive eye-movement (I claim that rhyme is an acoustic not visual phenomenon; and that there is a tendency to indicate discontinuation by prolonging the last sounds in ordinary speech and blank verse too, as well as in music — where no rhyme is involved).

Keywords: rhyme, resonance, phonetic coding, psychological reality, Tennyson, Baudelaire

I use the term "psychological reality" in this sense: the extent to which the constructs of linguistic theory can be taken to have a basis in the human mind, i.e., to somehow be reflected in human cognitive structures. The main part of this article explores at length the human cognitive structures in which the

* Reuven Tsur was Professor at Tel Aviv University, Israel. Unfortunately he passed away while this issue of SMP was in preparation.

https://doi.org/10.12697/smp.2021.8.1.01

constructs of phonetic theory may be reflected. In the last section I discuss the psychological reality of sound patterns in Baudelaire's "Les Chats".[1] In physical terms, I define "resonant" as "tending to reinforce or prolong sounds, especially by synchronous vibration" (watch a video in which, on the piano, the overtones related to the notes pressed with the right hand resonate when a note with similar overtones is struck with the left hand: https://www.tau.ac.il/~tsurxx/ba-da-ga/Video@resonance%20(Converted).mov); we do not compare overtones item by item, just hear their synchronous vibration. In phonetic terms I define "resonant" as "where intense precategorical auditory information lingers in short-term memory" (see below).

Speech Perception and Reading – Vocal and Subvocal

What is the relationship between poetic language on the one hand, and acoustics and resonance on the other?[2] I shall argue that ordinary speech is designed to suppress resonance, whereas poetic language is designed to enhance it.

The experiencing of poetry, whether through listening to an oral performance, or through subvocal reading, begins with an act of speech perception. Speech is transmitted through a stream of acoustic information that, on the computer screen, has no resemblance to the perceived speech categories. According to the motor theory of speech perception (Liberman, Mattingly 1982), this precategorical acoustic information does not reflect the intended speech sounds, but the articulatory gestures that produce them. This information is recoded into speech categories by the listener *via* his own articulatory system, and then excluded from consciousness.

---

1 The major part of this article is an article commissioned by a journal of applied physics for a special issue on acoustics and resonance in musical instruments. I concentrated in it what I had written on acoustics and resonance in poetry. Eventually I added, for literary readers, the last section in which I apply this discussion to issues raised in three articles on Baudelaire's "Les Chats".

2 This article summarizes what I have expounded at length in my earlier publications in a specific area of my research: the role of resonance in the sound patterns of poetry. The comprehensive theory underlying it has been propounded at book length in Tsur (1992; Tsur and Gafni forthcoming). In addition, this article integrates much of my other publications, repurposing them for the present argument. The parts reporting the experiments on sound–shape symbolism and double-edgedness in speech perception are derived from Tsur and Gafni (2019a), with the kind permission of the editor of *Literary Universals*.

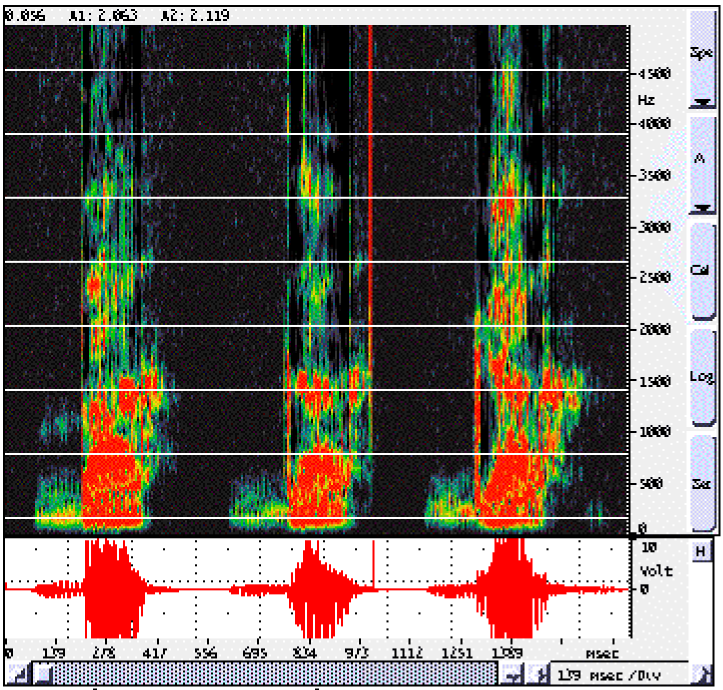
Recent brain research found that this process is somewhat more complex. Patricia K. Kuhl, Rey R. Ramírez, Alexis Bosseler, Jo-Fu Lotus Lin, and Toshiaki Imada (2014) investigated, using magnetoencephalography (MEG),

> motor brain activation, as well as auditory brain activation, during discrimination of native and nonnative syllables in infants at two ages that straddle the developmental transition from language-universal to language-specific speech perception. […] MEG data revealed that 7-mo-old infants activate auditory (superior temporal) as well as motor brain areas (Broca's area, cerebellum) in response to speech, and equivalently for native and nonnative syllables. However, in 11- and 12-mo-old infants, native speech activates auditory brain areas to a greater degree than nonnative, whereas nonnative speech activates motor brain areas to a greater degree than native speech.

My argument is based on the assumption that in certain circumstances the rich precategorical auditory information reverberates subliminally in short-term memory, and that certain well-known tasks and effects crucially depend on it, such as reading, versification and phonetic symbolism. First, as we shall see in the experiments with efficient and poor readers, during reading, the auditory information must subliminally reverberate in short-term memory, so as to render the phonetic material available while processing the text. Second, the repeated sounds in versification, such as in rhyme and alliteration, enhance each other's auditory information rendering them perceptually more salient and more memorable. Third, "phonetic symbolism" refers to a significant interaction between the sound and meaning of words. For reasons to be explained below, depending on the meaning, repeated nasals may be perceived as imitating some lowly-differentiated, continuous, low-pitch noise, as in "murmur", or may be perceived as expressive of some tender emotion, as in "meek and mild" generating, in a text stretch of some length, an emotionally laden atmosphere. Speech sounds have a wide range of sometimes conflicting meaning potentials; the meaning of the text picks out the relevant potential – then sound and meaning reinforce in each other their shared potentials. When the meaning has to do with sounds, it may sometimes amplify the auditory information so as to be perceived as resonating.

The key to my ensuing discussion is *phonetic coding.* Vowels can be uniquely identified by concentrations of overtones called "formants". Formants appear in a spectrogram one above the other, marked F1, F2, F3, etc. The upper window of Figure 1 presents the spectrograms of the syllables [ba], [da], and [ga].

Liberman et al. (1967) distinguish speech mode and nonspeech mode in aural perception. When we record sonar, we receive on the computer screen a pattern that is similar in shape to what we hear. This is the nonspeech mode. In the speech mode, the perceived speech category does not resemble the precategorical auditory information that conveyed it and was excluded from awareness. This is called "encodedness". Some speech sounds are more encoded, some less; that is, in some speech sounds less auditory information reaches consciousness, in some more. Thus, most people cannot discern which syllable is acoustically higher, /ba/, /da/, or /ga/; but can easily discern that /s/ is higher than /š/ and [i] is higher than [u].



ba da ga

Figure 1. Spectrogram of the syllables /ba/, /da/ and /ga/.

In certain artificial laboratory conditions, one can hear directly the precategorical auditory information. So, the reader may listen online to the sequence of syllables represented in Figure 2, and to the sequence of isolated second formant transitions from an unpublished demo tape by Terry Halwes ("Formants" are concentrations of overtones that uniquely identify vowels; "formant transitions" are the rapid change in frequency of a formant for a vowel immediately before or after a consonant, and give information about the vowel and the consonant simultaneously). See whether you can hear a gradual change between the steps, or a sudden change from /ba/ to /da/ to /ga/.

[Listen: https://www.tau.ac.il/~tsurxx/ba-da-ga/ba_da_ga.mp3] Halwes then isolates the second formant transition, that piece of sound which differs across the series, so as to make it possible to listen to just those sounds alone. [Listen: https://www.tau.ac.il/~tsurxx/ba-da-ga/Glide_and_whistles.mp3] One may discern two main perceptual differences between the two series. First, in the ba-da-ga series we hear no pitch differences, whereas in the chirp series we hear gradual pitch change, even though the latter series is excised from the former one. Secondly, most people who listen to that series of chirps report hearing what we would expect, judging from the appearance of the formant transition: upward glides, and falling whistles displaying a gradual change from one to the next; but not in the syllable series. Since these are hand-painted spectrograms, well-controlled experiments can be conducted. Participants who listen to the series of glides and whistles discriminate all the distances equally well. Participants who listen to pairs of consecutive stimuli from the ba-da-ga series, discern better the same distance at category boundaries than within categories (Mattingly et al. 1971). This is called "categorical perception". The perception of the syllable series illustrates the speech mode, of the chirp series – the non-speech mode.

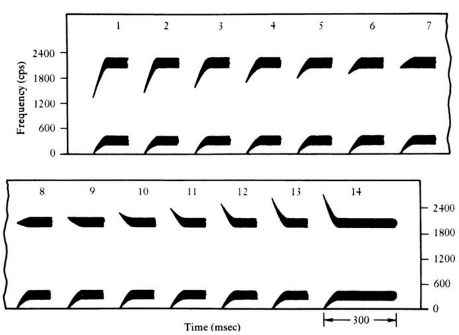


Figure 2. Hand-painted spectrograms of the syllables ba, da, ga. The ba–da–ga pitch continuum of F_2 is divided into 14 steps instead of three. The two parallel regions of black indicate regions of energy concentration, F_1 and F_2. Notice that the onset frequency of F_2 of da is higher than that of ba; and the onset frequency of F_2 of ga is higher than that of da. Only stimulus 14 represents its full duration. (Replayed in the accompanying sound file).

I have spoken of the "poetic mode of speech perception", where some of the precategorical auditory information is perceived behind the speech categories. In some poetic contexts this precategorical auditory information is perceived as more than usually "resonant". Let us further illustrate the foregoing generalizations through the following figures. Figure 3 is a natural speech spectrogram of the vowels [i] and [u].

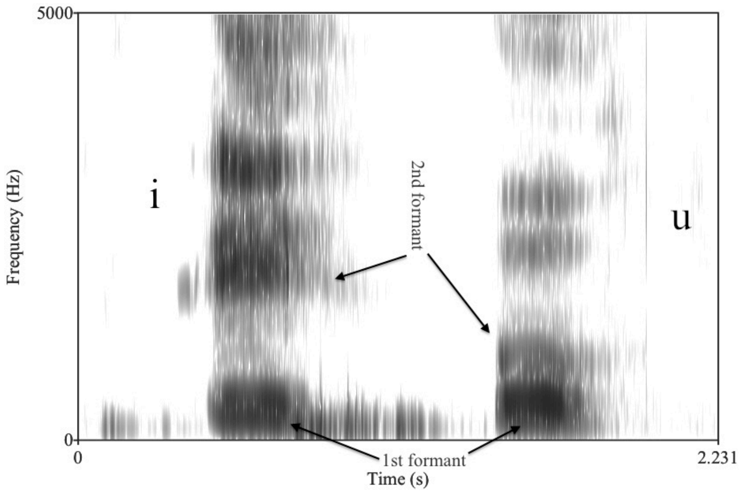


Figure 3. Spectrograms of the vowels [i] and [u] in natural speech

Formants, as we said, are overtones that uniquely define vowels. Formants reach up to the upper limit of the hearing range, but only the first two formants determine the vowels uniquely; the rest merely affect the quality of the perceived vowel. Simplified hand-painted spectrograms of vowels or syllables can be replayed through a computer.

In the hand-painted spectrograms of Figure 4, the parallel patterns reflect the first two formants of the vowels [i] and [u]. They are preceded by formant transitions. "Formant transitions" are the rapid change in frequency of a formant for a vowel immediately before or after a consonant, and give information about the vowel and the consonant simultaneously – this is called "parallel transmission". Note that the formant transition that conveys [d] (marked by a dotted circle) is different before [i] and [u]; and the parallel lines do not resemble the vowels they convey. The same pitch differences at arbitrarily-chosen points are heard as two-sound chords, not as unitary speech sounds. Briefly, as I said, there is no resemblance between the perceived speech sound and the sound wave that transmitted it. Figures 3–4 may explain why [i] is heard as higher than [u], cross-culturally, irrespective of the fundamental pitch. Likewise, Figure 5 can explain why we hear [s] higher than [š].

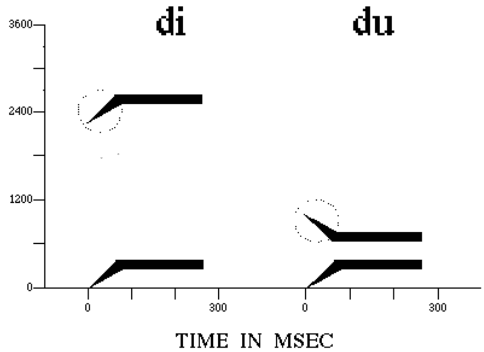


Figure 4. Simplified spectrographic patterns sufficient to produce the syllables [di] and [du].

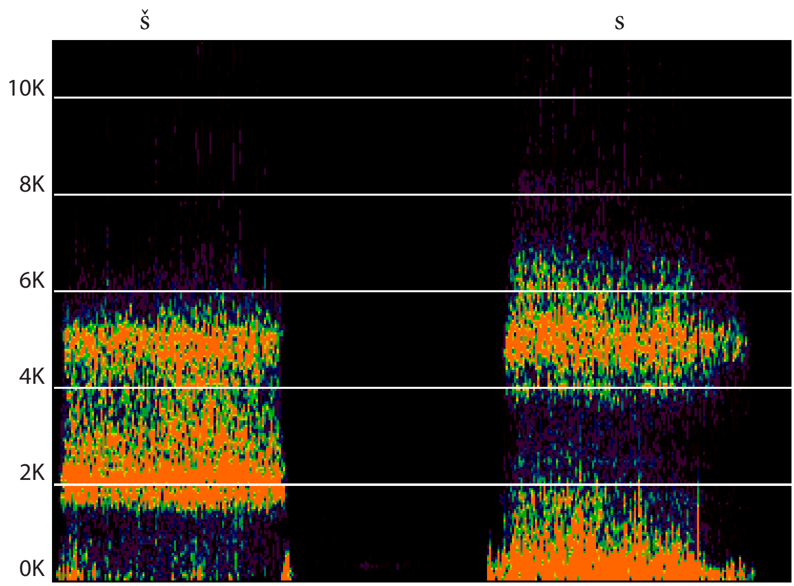


Figure 5. Sonograms of [š] and [s], indicating why [s] is somehow "higher".

I said above that the experiencing of poetry, whether through listening to an oral performance, or through subvocal reading, begins with an act of speech perception. Most persons hear an inner voice even when reading silently (Perrone-Bertolotti et al. 2012). Many people move their lips and tongues when reading. Even silent movement of the vocal tract may be perceived as "hearing", as the McGurk effect may suggest. Persons who see and hear a video showing the articulation of *i* but actually playing *u* hear something like *ü*, between the two. Preverbal infants who are shown two movies on two screens but hear only one or the other soundtrack, look at the screen whose soundtrack is played. They "know" which lip movements announce which sound series. Verdi is reported that when finishing an opera, he ended up with a sore throat: he subvocally moved his vocal tract with the roles he was writing, for all voice registers. Experiments with efficient and poor readers in elementary school revealed a cognitive mechanism based on phonetic coding and the reverberation of precategorical auditory information in short-term memory.

In several of my earlier publications (e.g., Tsur 1992; 2012; 2019), I summarised a series of experiments that revealed a cognitive mechanism supposed to be responsible for reading capabilities. Researchers at the Haskins Laboratories (e.g., Liberman, Mann 1981: 128–129; Brady et. al. 1983: 349–355; Mann 1984: 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; efficient readers, by contrast, seem to make 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 order to be able to count syllables or phonemes of words, the words must linger (that is, reverberate) in one's short-term memory. 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 precategorical acoustic information and, actually, enhance its memory traces. In the last section I will account for the psychological reality of rhyme in Baudelaire's "Les Chats" by this mechanism.

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. What in the nonaesthetic memory experiments is called acoustic *con**fusion**,* in an aesthetic context cooccurs with a coherent text, and may be perceived in the background as "harmonious ***fusion***", "musicality".

It is usually assumed that versification was developed in pre-literate societies as a memory aid: to verbatim memorize texts of cultural importance. Paradoxically, in this experiment rhyme confused rather than aided memory. *Confusion* occurs when the reverberating sound information of adjacent rhyme words mingle in short-term memory; rhyme *aids* memory when it organizes stretches of intervening, non-rhyming text: it chunks it into easily-perceptible chunks, and the "resonating" rhyme words enhance each other's memory traces enhancing, by the same token, the unity of the text.

Poor readers have recourse to some other kind of coding. Crowder and Wagner (1992: 228–230) summarize an experiment by Byrne and Shea (1979) 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. Crowder and Wagner insist: "The fact that poor readers seem to be using 'too much' meaningful processing does not imply that they are *better at* top-down processing than good readers. It is just that they may be so deficient in bottom-up processing they have no other recourse". In plain English, this means that they are guessing rather than processing.

Such precategorical acoustic information is, then, a prerequisite of efficient reading. We have recourse to it only subliminally. The sound patterns of poetry may amplify that information, so as to be perceived as vague "musicality". As we shall see, in some poetic contexts, such musicality may be further amplified, to be perceived as a resonant texture. That's why voiced plosives may be perceived differently in Excerpts 1 and 2 below.

The present argument makes it necessary to summarize a few distinctions of traditional phonetics. Consonants can be abrupt (–continuant): plosives (p, t, k, b, d, g) or affricates (as ts [in tsar], dž [in John or George], or pf [in German pfuj]); or, they may be continuous (+continuant), as nasals (m, n), liquids (l, r), glides (w [as in wield], or y [as in yield]); or fricatives (f, v, s, š [as in shield]). Continuous sounds may be periodic (nasals, liquids, glides) or aperiodic (fricatives). In periodic sounds, the same wave form is repeated indefinitely, while aperiodic sounds consist of streams of irregular sound stimuli. All vowels are continuous and periodic. Consonants can be unvoiced or voiced. All nasals, liquids and glides are voiced by default. Voiced plosives, fricatives and affricates (e.g. b, v, ž, and dž) consist, acoustically, of their unvoiced counterpart plus a stream of periodic voicing (cf. Tsur, Gafni 2019b).

Resonant and Nonresonant Speech Sounds

In the foregoing I described the cognitive mechanism underlying speech perception as well as vocal and subvocal reading. In what follows, I shall make four claims: that the same cognitive mechanism underlies the perception of resonance in poetry; that a resonant quality can be perceived in some poetic passages, in some it cannot; that the same speech sounds can be perceived as resonant in one context, but not in another; and that some persons are more sensitive to this difference than others.

I have said that ordinary speech is designed to suppress resonance, whereas poetic language is designed to enhance it. To substantiate this claim, I turn to a set of experiments on lateral inhibition (e.g., Crowder 1982a, 1982b), indicating that for biologically motivated reasons too, the same speech sounds are sometimes more and sometimes less reverberating. When speech sounds in proximity are very similar, they enhance each other's precategorical sound information in perception; when they are slightly similar, they inhibit each other; when they are not similar at all, there is no mutual effect. In other words, in some instances alliterating speech sounds may enhance reverberation in each other, in some – inhibit it. In ordinary speech, we typically focus on the final referent of the verbal message; according to Roman Jakobson (1960), the poetic function forces us to attend back to the verbal message. I wish to emphasize: "message" does not have here the vulgar sense of 'communicating some meaning', but is a total verbal structure encompassing meaning, syntax and sound. Jakobson's statement must be interpreted as that rhyme and alliteration serve to shift attention back to the sound structure, figurative language, parallelism etc., to the meaning; briefly, to make the reader linger on the whole complex message. But, in view of Crowder's experiments, sound patterns do much more: they enhance the reverberation of the rich precategorical auditory information; figurative language involves the reader in active problem solving in meaning.

Lateral inhibition has at least two important tasks in speech perception: it enhances phoneme oppositions and, at the same time, it prevents sound patterns from distracting attention from meaning. The sound patterns of poetry, on the contrary, serve to enhance the reverberation of precategorical auditory information, encompassing meaning.

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. Consider Excerpt 1:

1. And murmuring of innumerable bees.

The liquids and nasals /m/, /n/, /l/ and and /r/ and vowels in general are much less encoded than, e.g., the voiceless plosives /p/ and /k/; that is, in this heavily alliterating phrase some intense precategorical auditory information is perceptible. Furthermore, the meaning of the phrase foregrounds this precategorical auditory information, generating a very effective onomatopoeia. In this verse line, the syllable [mər] is repeated three times. The periodic sound waves perceived in the nasal, the liquids and the low back vowel are foregrounded by the meaning of "murmur". However, [mər] in "innumerable" is shorter than (that is, only slightly similar to) its first two tokens. Consequently, its reverberation tends to be inhibited and lost on the reader. Performers of this line tend, therefore, to prolong this syllable, so as to resonate with the other two tokens.

In the above line, we have both: "synchronous vibration" of precategorical auditory information; that is, much of the rich precategorical auditory information can be perceived in them, and in certain circumstances may be perceived as highly resonant.

Voiceless plosives, by contrast, are highly encoded, that is, no or little auditory information is perceptible in them. They are compact, "opaque". Voiced plosives consist of their unvoiced counterparts plus continuous, periodic voicing. They are ambiguous: in some contexts, voicing may bestow a massive presence on the plosive, in some – resonance.

I have said that voiced plosives are ambiguous: in some contexts, voicing may bestow a massive presence on the plosive, in some – resonance. The former is typically perceived as relatively hard, the latter as relatively soft, tender. A striking example of our third claim above (that the same speech sounds can be perceived as resonant in one context, but not in another) is provided by Iván Fónagy (1961), who in his classical paper, "Communication in Poetry", explored in poems by four poets in three languages the speech sounds that occur in tender and aggressive poems with more than regular frequency. He found that /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 exactly 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 (7.93:10.11), whereas for Hugo it has a basically tender quality (7.09:5.76) – again, in almost the same reverse proportion. If one decides that anything goes, one is exempt from bothering with it. If, however, one believes that this is significant, one should try to find some explanation that is consistent with what we know about mental processes in general, and speech perception in particular. The reason seems to be this. 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, Verlaine and Hugo applied the same cognitive mechanism to these voiced stops, but with a reverse focus.

Let us consider a minimal pair: Tennyson's line, with John Crowe Ransom's re-writing exercise.

1. And murmuring of innumerable bees
2. And murdering of innumerable beeves

There is general consensus that Excerpt 1 is perceived differently from Excerpt 2. Excerpt 1 is perceived by many readers as particularly resonant. Obviously, this is due to the repeated nasals [m] and [n] and the liquids [r] and [l], that are continuous, periodic, and lowly encoded. In Excerpt 2, however, "the euphony is destroyed […] we lose the echoic effect" (Abrams, Harpham 2009). This is not only because in 1 there is one nasal more than in 2, but also because in 1 the meaning foregrounds the resonant quality of voicing, so that the nasals and the liquids have a fuller, richer, more resonant body. There is an intuition that even [b] is perceived as more resonant in Excerpt 1 than in Excerpt 2. To put it bluntly, in Excerpt 2 it sounds hard and compact as plosives are supposed to sound; in Excerpt 1 it drifts slightly toward the more resonant quality of [m]. Here the question arises whether this difference has psychological reality, or this intuition results from mere transfer from the resounding meaning to the sound patterns. It cannot be tested explicitly by a stimulus–response questionnaire, because one cannot know whether participants respond to the meaning or the sound patterns of the line. Thus, a "less resonant" response in 2 could be due to the perception of [b] or merely to the meaning. So, less direct modes of testing are required, which I will provide later.

Empirical Evidence

In an experiment reported in Gafni, Tsur (2019; summarized in Tsur, Gafni 2019a) we asked participants to read out aloud pairs of consonant-vowel sequences (e.g. *ma-ba*) and compare the consonants on various bipolar perceptual scales (e.g. whether *m* sounds smoother and *b* sounds jerkier, or vice versa). Our experiment was designed specifically to test the hypothesis that voiced plosives are double-edged by contrasting them with voiceless plosives, on the one hand, and with nasals, on the other hand. Our hypothesis was largely supported. On the smoothness scale, we obtained a three-way contrast that was statistically significant: nasals were perceived smoother than voiced plosives that, in turn, were perceived as smoother than voiceless plosives. On other scales, we obtained only partial contrasts (possibly due to lack of statistical power): nasals were perceived as having fuzzier boundaries than voiced plosives that, in turn, were perceived as having fuzzier boundaries than voiceless plosives. However, the latter contrast was only 'near significant'. In addition, we received two one-sided contrasts: first, voiced plosives were perceived as harder than nasals, but voiceless plosives were not perceived as harder than voiced plosives. Second, voiced plosives were perceived as thicker than voiceless plosives, but nasals were not perceived as thicker than voiced plosives. Whether these imperfect results reflect the true state-of-affairs or not, they clearly demonstrate that voiced plosives are perceptually ambiguous: in certain contexts, they contrast with nasals and, in others, with voiceless plosives.

More interestingly for present purposes, our experiment also yielded a puzzling finding. Voiced plosives were perceived as having more resonance than voiceless plosives, though the result was only near significant. This was well expected. However, contrary to expectation, voiced plosives were also perceived as having more resonance than nasals (the result fell short of statistical significance after correction for multiple comparisons). A simple explanation for this unexpected result is that the task was not clear to the participants. As a matter of fact, two participants commented that they had trouble evaluating resonance, and, in general, it became clear to us that many people don't know what resonance means. However, there is also a possibility that the results are genuine, namely that voicing couples with plosion and endows voiced plosives with a resonating quality, which can be perceived, at least in controlled experiments, out of context. This hypothesis is supported by experimental evidence that voiced plosives are perceived as larger than voiceless plosives cross-linguistically (Shinohara, Kawahara 2016).

To make sure that experimental results do not depend on participants' understanding of a complicated task, one needs a more indirect experimental task, where simpler decisions must be made, and participants do not suspect that they are tested on the ambiguity of voiced plosives, that is, whether voiced plosives can be perceived as similar to voiceless plosives or, alternatively, to nasals. We, therefore, turn to a series of experiments in phonetic symbolism, repurposed to the issue in hand by Tsur, Gafni 2019a. The earliest of these experiments contained a forced-choice task, the later ones an interference task.

One of the most discussed examples of phonetic symbolism is that of sound–shape symbolism. Back in the nineteen twenties, Köhler (1929) took two nonsense words, *takete* and *baluma*, and asked people to match them with two nonsense figures, one with angular edges and one with rounded edges (see Figure 6). An overwhelming majority of respondents matched *takete* with the angular shape, *baluma* with the rounded shape.[3] Since Köhler's study, there were many replications of this effect, including a study by Ramachandran and Hubbard (2001), which used the nonsense words *bouba* and *kiki*. So, this effect came to be called the 'bouba/kiki' effect.

Figure 6. Shapes used in 'bouba/kiki' experiments (from Ramachandran, Hubbard 2001).

The Bouba/kiki effect has been demonstrated cross-culturally, even with Himba participants of Northern Namibia who had little exposure to Western cultural and environmental influences, and who do not use a written language (Bremner et al. 2013). Thus, the tendency to associate certain shapes with certain speech sounds is a cross-cultural phenomenon.

In all these experiments, the performance involved explicit decision-making. In what follows we shall present three more recent studies that attack the

3 In the second edition of his book, Köhler (1947) changed *baluma* to *maluma*.

problem at a pre-semantic, pre-conscious and pre-verbal level, respectively. There are quite a few studies that utilize interference tasks in exploring phonetic symbolism, that is, a demonstration of cognitive interference where a delay in the reaction time of a task occurs due to a mismatch in stimuli. None of them, however, set out to explore this issue of aspect switching between resounding and non-resounding voiced plosives. On the contrary, they are concerned with establishing phonetic symbolism as a consistent, involuntary, pre-rational phenomenon. But the various studies used the plosives in different experimental designs. Taken together, they provide strong evidence that voiced plosives can be perceived as resonant when contrasted to voiceless plosives, and as an abrupt plosive when contrasted to continuants.

In a study by Westbury (2005), participants performed a pre-lexical decision task on words and nonwords presented within curvy and spiky frames (word/nonword). Within each lexical category (word and nonword), some stimuli contained only stop consonants (e.g. toad and kide), some stimuli contained only continuous (e.g. moon and lole), and some stimuli that contained both stops and continuous sounds (e.g. flag and nuck). It was found that responses were faster for congruent shape-string pairs (continuous sounds in curvy shapes, plosives in spiky shapes) than for incongruent pairs. Using a masking technique, Hung, Styles and Hsieh (2017) showed that the mapping for the *bouba/kiki* effect occurs prior to conscious awareness of the visual stimuli. Under continuous flash suppression, congruent stimuli (e.g. "kiki" inside a spiky shape) broke through to conscious awareness faster than incongruent stimuli. This suggests that "a word [can] sound like a shape before you have seen it." This was true even for participants who were trained to pair unfamiliar letters with auditory word forms. These results show that the effect was driven by the phonology, not the visual features, of the letters.

In another study Ozturk, Krehm and Vouloumanos (2013) presented 4-month-old infants with pairs of shapes and auditory stimuli. They found that the infants looked longer at the screen during trials with incongruent pairs (i.e. 'bubu' with an angular shape or 'kiki' with a curvy shape) than during trials with congruent pairs (i.e. 'bubu' with a curvy shape or 'kiki' with an angular shape). This finding, together with cross-cultural evidence, suggests that at least some aspect of sound–shape symbolism is pre-linguistic, perhaps even innate. But which aspect exactly? I claim that what is innate is not the specific symbolic relation per se, but rather the propensity creatively to extract, contrast and compare abstract features from sensory stimuli.

Westbury put the voiced plosives in one bin with voiceless plosives, contrasting them to continuants. The other two interference tasks contrasted the voiced plosive of "bouba" with the voiceless plosives of "kiki" (Köhler had put the voiced plosive /b/ in one word with a liquid and a nasal ("baluma"), contrasting it with a word consisting of voiceless plosives). Thus, not only the participants, but even the experimenters were not aware that they were demonstrating the perceptual ambiguity of voiced plosives.

Sensitivity to Resonating Speech Sounds

Finally, I assume (Tsur 2012) that not all persons are equally sensitive to the distinction between resonant and nonresonant speech sounds. Obviously, efficient readers, who make good use of phonetic coding in short-term memory, are more sensitive than poor readers, who make inefficient use of phonetic coding. But, paradoxically, there is also a group of particularly efficient readers who are less sensitive to this distinction: speed readers. This assumption is supported by recent brain research. Most people may note that, when reading, an inner voice enunciates the words. Perrone-Bertolotti et al. (2012) explored the neural correlates of this "inner voice" in silent reading in four epileptic human patients recorded with intracranial electrodes in the auditory cortex for therapeutic purposes, and measured high-frequency (50–150 Hz) "gamma" activity (gamma brain waves are the fastest brainwave frequency with the smallest amplitude). Their findings are compatible with the assumption that relying on phonetic coding facilitates reading. But they also provide information about expert readers who tend to ignore the sound structure of poetry. In harmony with our foregoing discussion, they suggest that reading might rely more on phonological processes as texts become more difficult to read. Phonological activation would be more active as linguistic complexity increases, or in non-proficient readers, and triggered by top-down attentional processes.

> Auditory verbal imagery would thus facilitate verbal working memory, as suggested by several authors (Smith et al. 1995; Sato et al. 2004), to process the sentence as a whole, and not as a collection of unrelated pieces. Simultaneously, the activation of phonological representations in TVA [Temporal voice-selective area] would produce the vivid experience of the inner voice. (2012: 17560)

Perrone-Bertolotti et al.'s results relate to a long-standing debate as to whether expert readers automatically access phonological representation when reading since it would be difficult to think of a phonological representation that would not include an auditory imagery component:

> One neural possible explanation is that learning to read might strengthen the connectivity between visual and auditory areas (Booth et al. 2008; Richardson et al. 2011) based on Hebbian plasticity: both regions would be repeatedly coactivated because of repeated associations between visual and auditory inputs during the learning period (the written word and the auditory percept of one's own voice while reading overtly). With practice, this connectivity would allow for a direct activation of the auditory cortex by visual inputs through the visual cortex, in the absence of overt speech, very much like an automatic stimulus-response association. (2012: 17560)

Note that we are confronted with a mess of inconsistent terminology. In our foregoing discussion we spoke of "efficient" readers (as opposed to "poor" readers), who rely on phonetic coding. Perrone-Bertolotti et al. speak of adult "non-proficient" readers who rely on phonological processes. There may be a difference of standards for young children and adults. These adult "non-proficient" readers, however, are contrasted here to highly *expert* readers. We should, perhaps, speak of a scale of three categories of readers, labeled with descriptive rather than evaluative labels: readers who rely mainly on semantic coding, readers who rely on phonetic and phonological coding, and speed readers. Each later category is more efficient in reading than the preceding one. Readers of the first of these categories make little use of the auditory information. Readers of the other two categories make ample use of auditory information, but at different speeds. In speed readers, the relation between the visual grapheme and the auditory phoneme is over-practiced and becomes automatic, a quasi-synaesthetic relationship, in which the visual information triggers the phonological information.

Only a reading in which precategorical auditory information lingers for some time in short-term memory yields access to the musical dimension of poetry. Both poor readers and expert readers tend to ignore the sound dimension in reading, for opposite reasons: inefficient use of phonetic coding, and having over-practiced the phonetic coding, respectively. Perhaps, some speed readers can switch at will between reading styles: between dwelling on phonetic coding, and speed reading, according to the purpose of reading.

The Psychological Reality of Rhyme in Baudelaire's "Les Chats"

This last section discusses the psychological reality of rhyme in Baudelaire's "Les Chats". Far in its background we find Roman Jakobson and Claude Lévi-Strauss' article on this poem, who point out a huge number of parallelisms and oppositions on a large number of levels of this poem. This section was triggered by two later articles that raise, in one way or other, the issue of psychological reality regarding Jakobson and Lévi-Strauss's argument ([1962] 1987): Michael Riffaterre's on theoretical grounds (1966); and Fechino, Jacobs and Lüdtke's (2020) by conducting empirical tests. I will discuss only their comments on rhyme.

Neither Riffaterre, nor Fechino, Jacobs and Lüdtke use the explicit term "psychological reality". The former uses "objective" terms: "Can we not suppose, on the contrary, that the poem may contain certain structures that play no part in its function and effect as a literary work of art?" (2020: 202). In saying "play no part in its function and effect as a literary work of art", he may have meant "effect on the reader",[4] that is, "somehow reflected in human cognitive structures"; also, "the structures described do not explain what establishes contact between poetry and reader" (2020: 213). In fact, his essay explores the limits of the psychological reality of poetic devices. The latter conduct eye-tracking experiments to establish how the human mind processes the various patterns pointed out by Jakobson and Lévi-Strauss. While I enthusiastically welcome such attempts, I take exception to their interpretation of their findings on the sound patterns level.

The argument put forward here evolved during decades. When I first read Jakobson and Lévi-Strauss' article, I had an uneasy feeling that all those patterns pointed out by them cannot be experienced in a reading. Later, Riffaterre's article suggested an explanation to my intuition: not all linguistic patterns have aesthetic significance. Luckily, just when I was writing my article on acoustics and resonance, I ran into Fechino, Jacobs and Lüdtke's article. This coincidence induced me to make my own suggestions how the human cognitive system works in Baudelaire's rhymes.

Translation from Jakobson and Levy-Strauss' article:

3. Les amoureux fervents et les savants austères
Aiment également, dans leur mûre saison,
Les chats puissants et doux, orgueil de la maison,
Qui comme eux sont frileux et comme eux sédentaires.

Amis de la science et de la volupté
Ils cherchent le silence et l'horreur des ténèbres;
L'Erèbe les eût pris pour ses coursiers funèbres,
S'ils pouvaient au servage incliner leur fierté.

---

[4] I do not conceive of effects as of static features, out there, in a poem; I believe with Wellek and Warren (1949: 151) that a poem may be only "a *potential* cause of experience", that may or may not be realized in the various readings.

Ils prennent en songeant les nobles attitudes
Des grands sphinx allongés au fond des solitudes,
Qui semblent s'endormir dans un rêve sans fin;

Leurs reins féconds sont pleins d'étincelles magiques,
Et des parcelles d'or, ainsi qu'un sable fin,
Etoilent vaguement leurs prunelles mystiques.

Fervent lovers and austere scholars
Love equally, in their ripe season,
Powerful and gentle cats, the pride of the house,
Who like them are sensitive to cold and like them sedentary.

Friends of learning and of voluptuousness,
They seek silence and the horror of the shadows;
Erebus would have taken them as his gloomy coursers,
If they were able to incline their pride to servitude.

They assume in dozing the majestic poses
Of grand sphinxes reclining in the depth of solitudes
Who seem to be asleep in a dream without end;

Their fertile loins are full of magic sparks,
And particles of gold, like fine grains of sand,
Vaguely fleck their mystic pupils with stars

I was very much impressed by Fechino, Jacobs and Lüdtke's article "Following in Jakobson and Lévi-Strauss' footsteps: A neurocognitive poetics investigation of eye movements during the reading of Baudelaire's 'Les Chats'", but I have a few reservations too. Fechino, Jacobs and Lüdtke tested the psychological reality of the linguistic patterns pointed out by Jakobson and Lévi-Strauss, by tracking subjects' eye movements while reading Baudelaire's poem, then while re-reading it, both in a verse condition and in a prose condition (eliminating lineation).

> The significant interaction between Visual Presentation and Verse-Last Words and the results of separate analyses for verse and prose showed that the differences in the processing of rhyme words were only observed in the original verse form. In this verse condition, rhyme words dwelled upon longer, an effect visible in all duration-based eye tracking measures. In the prose condition, rhyme

> words were presented at all possible line positions […], but never occurred at the final position. Moreover, the words belonging to one rhyme pair did not occur at the same position in a line. Arranged this way, no significant differences in duration-based eye tracking measures were observed. Neither initial processing nor total reading time differed significantly between rhyme words and all other words. We therefore assume that readers identified rhyme pairs basically via regressive eye movements since the main effect for Verse-Last Words was significant also for rereading. (2020: 12–13)

This is an impressive, rigorous experiment. These researchers did not collect judgment, but tracked participants' eye-movements, of which the participants themselves were unaware. The problem is that rigorous experiment is one thing, its interpretation another. This needs a hypothesis. Reading poetry is a complex activity, involving visual, semantic, acoustic and phonetic processes. The authors derive their hypothesis from visual processes. In my view, poetic effects depend more on what we hear than what we see on the page, and working memory functions as an echo box; therefore, one may account for longer fixation time on rhyme words by "regressive eye movements" only if one has evidence for it in eye-tracking. Otherwise, one must look for acoustic and phonetic reasons. (The authors state that they *assume* "regressive eye movements", not that they rely on the findings of eye-tracking). My point of departure is "what our ear tells our mind". I believe that we discover sound similarity not through item-by-item comparison, but through resonance, that is, the reinforcement or prolongation of sound by the synchronous vibration of a neighbouring object. Jakobson and Lévi-Strauss point out in line 12 the internal rhyme "reins~pleins". For some reason they ignore that this line is embraced between lines 11 and 13, that rhyme on the homonymous words "fin". This close fourfold repetition of a nasal vowel generates an intensive texture of resonance, intensive musicality, pervading the whole "area". (Nasal vowels are exceptionally resonant, because in them, the breath resonates both in the oral cavity and the nose). Likewise, there is a brilliant sporadic rhyme (cf. Tartakovsky 2021) in lines 5–6: "science~silence", and another internal rhyme: "fervents~savant". Fechino, Jacobs and Lüdtke provide a table where they underline the rhyme words. These three pairs of sporadic rhyme are not underlined, nor mentioned in the article. So, we cannot know whether they found what they regard as indications of "regressive eye movements" in these internal rhyme words. Apparently, they did not. But the hypothesis of "regressive eye movements" ought to apply to them just as to formal rhymes.

The authors do, indeed, quote from my work one sentence on resonance: "Tsur (2002) stated that lines can be perceived as perceptual wholes (gestalts), if they can be contained in working memory, which functions in the acoustic mode like an echo box"; but this acoustic mode does not really influence their interpretation of their results. For me, the manipulation of duration is a problem-solving device, not merely a function of eye movements, both in speech and music. Longer duration, that is, slowing down at the end of a musical unit, suggests, according to Leonard B. Meyer, that no further progression is to be expected. This is the case in ordinary speech too, not only in end rhyme. In ordinary speech, we usually cue sentence ending by redundant acoustic cues: pause, terminal intonation contour and slowing down the last word or speech sounds. Since in verse the perceptual unity of the line crucially depends on its closure, these cues may be more emphatic in it than in ordinary speech. Since the eye-tracking technique gives information only on duration, not the other cues, we must take comparison between verse-final and prose-final words with a grain of salt. The white visual space around the printed verse gives instructions that in vocal performance one must have recourse to those phonetic cues for discontinuation.

This is the case, more emphatically, with blank verse and free verse, indicating line ending. We expose ourselves for a longer time to line-final words in blank verse and free verse too, where no rhyme is involved. The reason is that prolongation indicates line ending directly, rather than that it takes time to go back to compare the rhyme word to an earlier rhyme fellow. In patterns of sound repetition (rhyme and alliteration), the similarity effect is established not by "regressive eye movements", but by the resonance of similar sound patterns that mingle spontaneously, enhancing each other in perception. When the distance between the repeated sound patterns exceeds the span of working memory, the sound pattern may have intellectual appeal but no psychological reality. Riffaterre comments: "Jakobson and Lévi-Strauss, [p. 191] see a paronomasia – to my mind very far-fetched – linking *fervent* and *frileux*" (1966: 219). There is here certainly an opposition between heat and cold. The question is whether the two words are close enough in the poem to resonate together in short-term memory.

Thus, for instance, Jakobson and Lévi-Strauss write: "The two predicates, the first and last in the sonnet, are the only ones accompanied by adverbs, both of them derived from adjectives and linked to one another by a deep rhyme: $_{2}$*Aiment également* – $_{14}$*É*toi*lent* v*aguement*" ([1962] 1987: 184). The mid-rhyme in the second and last line of the sonnet is a conspicuous instance of a rhyme-pair that exceeds the span of resonance (but not the span of "regressive eye movements") and has, therefore, no psychological reality, only intellectual appeal, if at all. Also consider the phrase "both of them derived from adjectives and linked to one another by a deep rhyme". Intuitively, such fine-grained description may exceed the reader's limited channel capacity. More likely, he would respond to this rhyme (if its members were closer together) on a higher level of abstraction. Jakobson's distinction elsewhere, "grammatical and anti-grammatical rhyme" might be more relevant here (Jakobson, Halle 1956: 82). The former consists of words belonging to the same word class, or of similar suffixes; the latter consists of words belonging to different word classes. On this level, the reader must make only a same~different judgement, without going into further distinctions. Rhyme words consist of similar sounds and dissimilar meanings; in anti-grammatical rhyme, meaning is more different than in grammatical rhyme. Rather than "deep rhyme" (whatever it means), Wimsatt (1954: 160) calls grammatical rhymes "tame rhymes".

One of the differentia of poetic language is that one may find in it an intricate crisscross of relationships between its subtle elements. Some critics call it multivalence, or multiple relationship. It is hard to determine the legitimate limits of such crisscrossing. One of the beauties of Riffaterre's essay is that he elegantly detects how Jakobson and Lévi-Strauss transgress those limits. I cannot do here justice to Riffaterre's long and subtle essay; I will illustrate his conception by one illuminating example. "Jakobson and Lévi-Strauss take literally the technical meaning of *feminine* as used in metrics and grammar and endow the formal feminine categories with esthetic and even ethical values. They are trying to prove a sexual ambiguity in the poem, the motif of the androgyne, and they find some evidence in the 'paradoxical choice of feminine substantives [for] masculine rhymes' [p. 197]" (1966: 209). I would say that "feminine" is part of the metalanguage, not of the poem.

However, I disagree with him on two rhymes. On the internal rhyme *eux–frileux*, Riffaterre writes: "In context the difference outweighs the similarities": such an internal rhyme is obvious, "because identical stresses 'confirm'" it. Nevertheless, "any responsion to the rhyme in line 4 still appears purely theoretical" (1966: 207). He accounts for this by a longish argument based on the interaction of phrase structure, meter, and simile structure. In my mind, the explanation is simpler. The stresses on pronouns and content words are not "identical". There is some evidence that the sound structure of function words tends to have less weight in the sound patterns of poetry. If one wishes to hear *eux* as rhyming, one must slightly prolong it.

On the other rhyme, Riffaterre says: "a natural reading of line 12 will have to take into account the tight unity of *leurs reins féconds*, which demands a pause after *féconds*, the normal caesura disappearing almost because *pleins* cannot be severed from *d'étincelles*; *pleins* is enclitic, which practically cancels out the rhyme" (1966: 207). We agree on the facts, but disagree on their interpretation. While Riffaterre claims that the phrase structure "practically cancels out the rhyme", I argue that the phrase structure "*backgrounds* the internal rhyme". There is no reason on earth why should one stop after *pleins* in order to hear the resonance of the sound pattern in the background. If one wishes to foreground it, one may slightly isolate it without stopping, by prolonging the *n*. We are up against convergent and divergent style. In the former, language patterns converge with versification patterns, in the latter they diverge. This internal rhyme of Baudelaire's is mildly divergent. In the second stanza of his "Correspondances", sound repetition is exceptionally rich and exceptionally divergent.

In English poetry, Milton is one of the most divergent, but also of the most musical poets. Tennyson wrote in his "Milton": "O mighty-mouth'd inventor of harmonies, / […] God-gifted organ-voice of England". Yet, Milton's poetry contains an unusual number of metric deviations and run-on lines. In the first 165 lines of *Paradise Lost* there are only three regular iambic pentameter lines. Many critics are baffled by this. This paradox of exceptional musicality and exceptional divergence from versification can be settled by pointing out several aspects, of which I shall mention three. First, unconstrained deviation from versification leads to chaos; but deviation while observing certain constraints yields a weak gestalt, divergent style, characterized by an emotional quality and musicality. One of those constraints is effective closure of the line, generated by several acoustic cues, among them by slowing down the last speech sounds. In "Les Chats", line endings are well articulated by syntax, but this specific internal rhyme is somewhat divergent. In many of Milton's lines, line endings are threatened by run-on lines, and sound repetitions are rich and highly divergent. In such run-on lines, experienced readers secure the perception of line endings without arresting, by slowing down the last speech sounds, and some additional acoustic cues (see below).

Second, in divergent style, alliterations are backgrounded. In the first few lines of *Paradise Lost*, the exceptionally rich alliteration escapes the attention of many readers, it just sounds very musical. In the second quatrain of Baudelaire's "Correspondances", alliteration is exceptionally rich and diffuse; yet, Henri Peyre, professor of French poetry at Yale, said that there are no virtuoso sound effects in the octet, acknowledging, nevertheless, its overwhelming effect. The backgrounded sound repetitions escaped his attention. Third, gestalt experiments indicate that colour interaction in paintings and overtone interaction in music is boosted within strong gestalt boundaries and inhibited across them; across weak gestalt boundaries, by contrast, it is boosted. The effect of rhyme and alliteration is achieved through overtone interaction in patterns of similar speech sounds that are backgrounded and boosted across the run-on line endings – hence its mysterious musicality. A single complex sentence runs through the second quatrain of "Correspondances", blurring the line endings. The main clause begins with an adverbial in its first line, and ends with the subject and predicate in the last line, generating high predictive tension. In "Leurs *reins* féconds sont *pleins* d'étincelles magiques", by contrast, the end of the clause coincides with the end of the line, effectively closing it and boosting overtone interaction between similar words. The deviation of the sound pattern from the phrase structure in "*reins* féconds" weakens the gestalt; by the same token, it renders the sound pattern more elusive. What is more, one must lengthen the *n* of "pleins" in order to articulate the caesura without stopping.

One important difference between behaviourism and cognitive studies is that the former is interested in the stimulus that goes into the system, and the response that comes out; it is not interested in the mediating structures in that black box, the skull. Cognitive studies, by contrast, concentrate on the mediating structures and the information processing that connect between stimulus and response. Eye-tracking techniques are powerful tools to explore the mediating structures and processes in the head. But to learn about what our ear tells our mind, we must interpret our eye-tracking findings within hypotheses drawn from acoustic and phonetic research. In the present case, longer fixation indicates that we slow down on line-final words even when reading sub-vocally; that is, that we *hear* what we read in poetry even when we do not vocalize. Speed readers, I suppose, do not fixate longer on line-final words, and do not – I suppose – *hear* line endings.

The same I profess about the authors' discussion of "visual space" at the end of their article.

> In poetry, space management is a central feature deliberately shaped by the poet. As highlighted by Derrida (1972), "spacing" is an active, productive characteristic of space which could become a medium of communication. While West-Pavlov, the main advocate of space theory, did not consider the reader in its theory, other scholars, such as La Charité (1987) considered both readers and eye movements. He proposed that space management could guide gaze path to build sentence semantics. Indeed, the prose visual presentation does not take into account space managing. (Fechino, Jacobs, Lüdtke 2020: 14)

I argue, by contrast, that "space management" does not merely "guide gaze path", but also gives a set of phonetic instructions to articulate line ending. Considering that we are dealing with the sound effects of poetry, their article gives disproportionately little information about what our ear tells our mind, as compared to what our eye tells our mind.

I am quoting from my article in *Style* "Free Verse, Enjambment, Irony":

> In prose, the lines run from one margin of the page to the other; in poetry, it is the poet who decides where the line ends. Fraser (1970: 29) writes about blank verse: "... we might often be uncertain (particularly when the sense is run on from line to line, without punctuational pauses at the end of the line) how the lines divide. This is what Dr. Johnson meant when he said that English blank verse is often verse for the eye".[5] If this were true, it would apply even more to free verse. Indeed, the surrounding empty space that indicates line endings in printed verse is not available in vocal performance. I claim, however, that just as white spaces break up the series of black marks on the paper into smaller perceptual units whose end may or may not coincide with the end of syntactic units, in aural perception, certain vocal devices may break up the text into versification units, and even indicate conflicts of versification and syntactic units. Indeed, Dr. Johnson does grant what Fraser is tacit on: there *are* "a few skilful and happy readers of Milton who enable their audience to perceive where the lines end or begin.
>
> In vocal performance there are other means to indicate line ending: first of all, "punctuational pauses"; but also intonation contour, and some more elusive cues, such as the lengthening of the last speech sounds or syllable, or over-articulation of the word boundaries, e.g., by inserting a stop release or a glottal stop where appropriate.[6] Such cues may act in conjunction – indicating unambiguous continuity or discontinuity; or in conflict – indicating continuity and discontinuity at the same time. (Tsur 2015: 36)

For the past twenty years or so, I have demonstrated how oral enjambment can be generated by inserting discontinuity electronically into continuous readings, without abolishing continuity.

In my article (Tsur 2015), I have shown how this works in enjambments in free verse. In our article "Enjambment – Irony, Wit, Emotion" in *Studia*

[5] "… there are only a few skilful and happy readers of Milton who enable their audience to perceive where the lines end or begin. 'Blank verse,' said an ingenious critick, 'seems to be verse only to the eye'. […] He that thinks himself capable of astonishing may write blank verse, but those that hope only to please must condescend to rhyme" (Johnson, "Milton", in *Lives of the Poets*). Dr. Johnson, as so often, seems to be ironical about that "ingenious critick"; Frazer takes him literally. My work explores the devices by which those "skilful and happy readers of Milton [...] enable their audience to perceive where the lines end or begin".

[6] A glottal stop is the sound we may insert in words beginning with a vowel, as in "I said 'an aim', not 'a name'".

Metrica et Poetica, Chen Gafni and I demonstrated this principle with reference to Milton's sonnets and blank verse (Tsur, Gafni 2018; included in Tsur, Gafni forthcoming). We found three recordings of Milton's sonnet "On his Blindness". We focused on the lines

4. who best
Bear his mild yoke, they serve him best. His state
Is kingly: thousands at his bidding speed

One performance preserved the continuity of the sentence "His state/ Is kingly", suppressing the line ending (listen https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9658). The other two performances preserved the line ending by observing a pause and a rising-falling intonation contour on "state" – suppressing sentence continuity. We electronically manipulated the duration and intonation of "his state" in the continuous performance, without inserting a pause. Lengthening the word "state" and changing its intonation contour suggested line ending, but the absence of pause suggested the continuity of the sentence. We also manipulated "thousands", to ensure a caesura after it, without a pause (listen to the genuine and doctored performances of "his state": https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9661; https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9662).

Milton's blank verse abounds in longish run-on passages. James Whaler found in his book (1956) *Counterpoint and Symbol: An Inquiry into the Rhythm of Milton's Epic Style* that quite a few of those run-on passages could be rearranged as non-enjambed passages, still preserving Milton's iambic pentameter (hence "counterpoint"). Whaler didn't quite know how to extract significance from this finding. But we dealt with one of his pairs of word-by-word identical passages in two ways.

5. 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.

(*Paradise Lost* IV. 917–923)

6. 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.

We submitted a print-out of the two versions to a panel of judges, who perceived the ironic contents as more ironical in the run-on version; we offered a gestaltist hypothesis to account for this. Leonard B. Meyer, who applies gestalt theory to music, accounts for the association of strong gestalts with 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). The opposite applies to weak shapes and emotional qualities. We explain how such psychological atmospheres interact with meanings in poetry. The divergent structure in Excerpt 5 seems to affect not only emotional qualities, but irony too, rendering it subtler. Meyer's formulation may account for this effect too. The ironist pretends to know nothing, not even that he is ironical. The "psychological atmosphere of certainty, security, and patent purpose" of strong gestalts subverts, therefore, the tone of elusive ignorance in irony.

In addition, we asked experienced readers to perform the two versions. In the run-on version, one could hear at the line endings continuity and discontinuity at the same time. Line discontinuity was indicated by word lengthening and some additional cues; sentence continuity was indicated mainly by the absence of a pause. Thus, the same words were longer when they were line-final in one version than when they occurred in midline in the other version. Briefly, we contrasted not a verse version and a prose version of the same text, but two literally identical verse versions that differed only in lineation (listen to the first one and the half lines in Excerpts 4, https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9663 and 5, https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9664). *Thee* is considerably longer in Excerpt 5, https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9665 than in Excerpt 6, https://ojs.utlib.ee/index.php/smp/article/view/smp.2018.5.2.01/9666. "Alone" ends a sentence in both versions; therefore there is no significant difference between them.

So, I would not be so sure in Fechino, Jacobs and Lüdtke's conclusion: "We therefore assume that readers identified rhyme pairs basically via regressive eye movements since the main effect for Verse-Last Words was significant also for rereading." One should not, therefore, necessarily assume "regressive eye movement" for "the main effect for Verse-Last Words", because the effect is inherent in the dynamics of the sound stream, even where no rhyme is involved, as in blank verse and free verse; and even where no words are involved at all, as in music. The need to foreground line ending would also explain why does fixation occur in re-reading too.

Moreover, even in rhymed verse we need not assume "regressive eye movements" (unless one gets explicit evidence from eye-tracking), because we discover sound similarity not through item-by-item comparison, but through resonance, that is, the reinforcement or prolongation of sound by the synchronous vibration of a neighbouring object. That is why the effect of sound patterns is *immediate*.

To Sum Up

We have explored poetic language and its relationship to acoustics and resonance, both in vocal recital and in subvocal reading. We started with the acoustic and motor mechanisms involved in speech perception. We have adduced evidence that some of those mechanisms are active even when no audible sound is produced. Most readers hear an inner voice when reading silently, some even move their lips and tongues when reading; the neural correlates of this have even been traced in the brain (Perrone-Bertolotti et al. 2012). In many of our publications, Chen Gafni and I have emphasized the acoustic ingredient in phonetic symbolism (only briefly mentioned here) and in accommodating the conflicting patterns of language and versification in a vocal performance. I conceive of speech sounds as of bundles of acoustic, articulatory and phonological features. Phonological features serve for arbitrary reference; emotive and sound-symbolic suggestions as well as rhythmical solutions are carried by the acoustic and articulatory features.

Based on the findings of speech research, I have assumed the "poetic mode of speech perception" (Tsur 1992), where some of the rich precategorical auditory information that transmitted speech (usually excluded from awareness) does reach consciousness. With reference to Excerpts 1–2 I have pointed out that in some contexts, owing to interaction with meaning, sound patterns are perceived as more than usually resonant. Most critics agree that in Excerpt 1, but not in 2, meaning exceptionally foregrounds in perception the periodic sound waves in the liquids and nasals, yielding stronger than usual reverberation. I claimed that in Excerpt 1 even the voicing ingredient of the repeated

[b] tends to be perceived as more reverberating than in Excerpt 2. In a stimulus–response experiment we found, contrary to our expectations, that our participants perceived voiced plosives as more resonant than nasals. While we were pleased to obtain experimental support for the possibility that in Excerpt 1 [b] is perceived as resonant, we did not rule out the possibility that some of our participants did not know what "resonant" means, or had difficulties with making resonance judgments. To our pleasant surprise, researchers who applied the Stroop effect to test the *bouba/kiki* effect provided, without knowing, straightforward evidence for the perceptual ambiguity of voiced plosives.

In the last section I discussed Jakobson and Lévi-Strauss' article on Baudelaire's "Les Chats", and two later articles that raise the issue of the psychological reality of their arguments. Though I accept generally, Riffaterre's criticism, I also have some reservations from it. Fechino, Jacobs and Lüdtke's eye-tracking technique is an efficient indirect way to find out what happens in that black box, the skull. But there are quite a few missing links to the sound effects of poetry. More importantly, they look for an explanation of the rhyme's effect in eye movements and not in acoustic and phonetic processes. I don't know enough about Lévi-Strauss and sounds. But Jakobson believed that we *hear* the sound similarities. Fechino, Jacobs and Lüdtke attribute to rhyme its effect by *assuming* regressive eye movements, do not rely on the findings of eye-tracking. When Riffaterre criticized Jakobson and Lévi-Strauss' paper saying that some of those sound similarities cannot have psychological reality, he did not mean that readers are incapable of those regressive eye movements, but that syntactic structure prevents the reader from directly experiencing the sounds of some of those internal rhymes.

Briefly, without applying the researcher's personal intuition to poetry, there is little chance that researchers would "hit" upon an adequate hypothesis to account for the poetic quality. Jakobson and Lévi-Strauss as well as Fechino et al. attempt to eliminate any personal bias, the former by pointing out an enormous number of linguistic parallelisms and oppositions, without asking whether they have psychological reality; the latter, by conducting objective eye-tracking experiments and performing "objective" statistics. By excluding their own personal experience, they looked for, and found, "scientific" support for far-fetched hypotheses. Riffaterre and I myself start with a personal bias, what our ear hears, and then look for support in the text. The difference between us is that Riffaterre acknowledges an either/or situation: the device either does, or does not have psychological reality; I distinguish two additional possibilities: the devices may be backgrounded or foregrounded, and I describe their respective conditions. Thus, some devices that have no psychological reality for Riffaterre, may be merely backgrounded for me.

References

Abrams, Meyer Howard; Harpham, Geoffrey 2014. *A Glossary of Literary Terms*. Andover, Hampshire: Cengage Learning.

Brady, Susan; Shankweiler, Donald; Mann, Virginia 1983. Speech perception and memory coding in relation to reading ability. In: *Journal of Experimental Child Psychology* 35(2), 345–367. https://doi.org/10.1016/0022-0965(83)90087-5

Bremner, Andrew J.; Caparos, Serge; Davidoff, Jules; de Fockert, Jan: Linnell, Karina J.; Spence, Charles 2013. "Bouba" and "Kiki" in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. In: *Cognition* 126(2), 165–172. https://doi.org/10.1016/j.cognition.2012.09.007

Charité, Virginia A. L. 1987. *Dynamics of Space: Mallarmé's Un Coup de dés jamais n'abolira le hasard*. Lexington, Kentucky: French Forum Publishers.

Crowder, Robert G.; Wagner, Richard K. 1992. *The Psychology of Reading*. New York: Oxford University Press.

Crowder, Robert G. 1982a. A Common Basis for Auditory Sensory Storage in Perception and Immediate Memory. In: *Perception & Psychophysics* 31(5), 477–483. https://doi.org/10.3758/BF03204857

Crowder, Robert G. 1982b. Disinhibition of Masking in Auditory Sensory Memory. In: *Memory & Cognition* 10(5), 424–433. https://doi.org/10.3758/BF03197644

Derrida, Jacques [1966] 1978. Structure, Sign, and Play in the Discourse of the Human Sciences. In his *Writing and Difference*. Translated by Alan Bass. Chicago: The University of Chicago Press, 278–293.

Fechino, Marion; Jacobs, Arthur M.; Lüdtke, Jana 2020. Following in Jakobson and Lévi-Strauss' footsteps: A neurocognitive poetics investigation of eye movements during the reading of Baudelaire's 'Les Chats'. In: *Journal of Eye Movement Research* 13(3), 4. https://doi.org/10.16910/jemr.13.3.4

Fónagy, Iván 1961. Communication in Poetry. In: *Word* 17(2), 194–218. https://doi.org/10.1080/00437956.1961.11659754

Fónagy, Iván; Magdics, Klara 1960. Speed of Utterance in Phrases of Different Lengths. In: *Language and Speech* 3(4), 173–192. https://doi.org/10.1177/002383096000300401

Fraser, George Sutherland 1970. *Metre, Rhyme and Free Verse.* London: Methuen.

Gafni, Chen; Tsur, Reuven 2019. Some Experimental Evidence for Sound-Emotion Interaction. In: *Scientific Study of Literature* 9(1), 53–71. https://doi.org/10.1075/ssol.19002.gaf

Halwes, Terry. Unpublished demo tape.

Hung, Shao-Min; Styles, Suzy J.; Hsieh, Po-Jang 2017. Can a Word Sound Like a Shape Before You Have Seen It? Sound-Shape Mapping Prior to Conscious Awareness. In: *Psychological Science* 28(3), 263–275. https://doi.org/10.1177/0956797616677313

Jakobson, Roman 1960. Closing Statement: Linguistics and Poetics. In: Sebeok, Thomas A. (ed.), *Style in Language*, 350–377. Cambridge, MA: The MIT Press.

Jakobson, Roman; Lévi-Strauss, Claude [1962] 1987. Baudelaire's "Les Chats"., In: Jakobson, Roman. *Language in Literature*. Cambridge, Mass., and London: The Belknap Press of Harvard University Press, 180–197.

Jakobson, Roman; Halle, Morris 1956. *Fundamentals of Language*. The Hague: Mouton.

Johnsson, Dr. Samuel. "Life of Milton" (1779, excerpt). https://www.poetryfoundation.org/articles/69382/from-lives-of-the-poets

Köhler, Wolfgang 1929. *Gestalt psychology*. New York: Liveright.

Köhler, Wolfgang 1947. *Gestalt psychology* (2nd ed.). New York: Liveright.

Kuhl, Patricia K.; Ramírez, Rey R.; Bosseler, Alexis; Lin, Jo-Fu Lotus; Imada, Toshiaki 2014. Infants' brain responses to speech suggest Analysis by Synthesis. *PNAS* 111(31): 11238–11245. https://doi.org/10.1073/pnas.1410963111

Liberman, Alvin M.; Cooper, Franklin S. Shankweiler, Donald P.; Studdert-Kennedy, Michael 1967. Perception of the Speech Code. In: *Psychological Review* 74(6), 431–461. https://doi.org/10.1037/h0020279

Liberman, Alvin M.; Mattingly, Ignatius G. 1985. The motor theory of speech perception revised. In: *Cognition* 21(1), l–36. https://doi.org/10.1016/0010-0277(85)90021-6

Liberman, Isabelle Y.; Mann, Virginia A. 1981. Should reading instruction and remediation vary with the sex of the child? In: *Status Report on Speech Research SR-65*. New Haven, CT: Haskins Laboratories, 125–143. http://www.haskins.yale.edu/sr/sr065/SR065_06.pdf

Mann, Virginia A. 1984. Reading Skill and Language Skill. In: *Developmental Review* 4(1), 1–15. https://doi.org/10.1016/0273-2297(84)90014-5

Mattingly, Ignatius G.; Liberman, Alvin M.; Syrdal, Ann K.; Halwes Terry 1971. Discrimination in Speech and Nonspeech Mode. In: *Cognitive Psychology* 2(2), 131–157. https://doi.org/10.1016/0010-0285(71)90006-5

Meyer, Leonard B. 1956. *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Ozturk, Ozge; Krehm, Madelaine; Vouloumanos, Athena 2013. Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. In: *Journal of Experimental Child Psychology* 114(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004

Perrone-Bertolotti, Marcela; Kujala, Jan; Vidal, Juan R.; Hamame, Carlos M.; Ossandon, Tomas; Bertrand, Olivier; Minotti, Lorella; Kahane, Philippe; Jerbi, Karim; Lachaux, Jean-Philippe 2012. How Silent Is Silent Reading? Intracerebral Evidence for Top-Down Activation of Temporal Voice Areas during Reading. In: *Journal of Neuroscience* 32(49), 17554–17562. https://doi.org/10.1523/JNEUROSCI.2982-12.2012

Ramachandran, Vilayanur S.; Hubbard, Edward M. 2001. Synaesthesia – A Window Into Perception, Thought and Language. In: *Journal of Consciousness Studies* 8(12), 3–34.

Richardson, Fiona M.; Seghier Mohamed L.; Leff Alex P.; Thomas, Michael S.; Price, Cathy J. 2011. Multiple Routes from Occipital to Temporal Cortices during Reading. In: *J Neurosci* 31(22), 8239–8247. https://doi.org/10.1523/JNEUROSCI.6519-10.2011

Riffaterre, Michael 1966. Describing poetic structures: Two approaches to Baudelaire's *les Chats*. In: *Yale French Studies* 36/37, 200–242. https://doi.org/10.2307/2930407

Sato, Marc; Baciu, Monica; Loevenbruck, Hélène; Schwartz, Jean-Luc; Cathiard, Marie-Agnès, Segebarth, Christoph; Abry, Christian 2004. Multistable representation of speech forms: a functional MRI study of verbal transformation. In: *Neuroimage* 23(3), 1143–1151. https://doi.org/10.1016/j.neuroimage.2004.07.055

Smith J. David; Wilson, Meg; Reisberg, Daniel 1995. The role of subvocalization in auditory imagery. In: *Neuropsychologia* 33(11), 1433–1454. https://doi.org/10.1016/0028-3932(95)00074-d

Tartakovsky, Roi 2021. *Surprised by Sound: Rhyme's Inner Workings*. Baton Rouge: Louisiana State University Press.

Tsur, Reuven 1992. *What Makes Sound Patterns Expressive: The Poetic Mode of Speech Perception*. Durham, N. C.: Duke University Press. https://doi.org/10.2307/j.ctv1131366

Tsur, Reuven 2012. *Playing by Ear and the Tip of the Tongue: Precategorial information in poetry*. (Linguistic Approaches to Literature, 14). Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/lal.14

Tsur, Reuven 2015. Issues in the Instrumental Study of Poetry. In: *Journal of literary theory*. 9, 1 (Empirical Methods in Literary Studies), 112–134. https://doi.org/10.1515/jlt-2015-0006

Tsur, Reuven 2015. Free Verse, Enjambment, Irony: A Case Study. In: *Style* 49, 35–46. https://doi.org/10.5325/style.49.1.0035

Tsur, Reuven 2019. Statistical versus Structural-Cognitive Approaches to Phonetic Symbolism: Two Case Studies. In: *Style* 53(3), 281–307. https://doi.org/10.5325/style.53.3.0281

Tsur, Reuven; Gafni, Chen 2018. Enjambment – Irony, Wit, Emotion. A Case Study Suggesting Wider Principles. In: *Studia Metrica et Poetica* 5(2), 7–28. https://doi.org/10.12697/smp.2018.5.2.01

Tsur, Reuven; Chen Gafni (forthcoming). *Sound–Emotion Relationship in Poetry.* Amsterdam: John Benjamins.

Tsur, Reuven; Gafni, Chen 2019a. Phonetic Symbolism: Double-Edgedness and Aspect-Switching. In: *Literary Universals Project*. https://literary-universals.uconn.edu/2019/07/20/phonetic-symbolism-double-edgedness-and-aspect-switching

Tsur, Reuven; Gafni, Chen 2019b. Methodological Issues in the Study of Phonetic Symbolism. In: *Scientific Study of Literature* 9(2), 195–229. https://doi.org/10.1075/ssol.19001.tsu

Wellek, René; Warren, Austin 1949. *Theory of Literature.* New York: Harcourt, Brace & Co.

West-Pavlov, Russell 2009. *Space in Theory: Kristeva, Foucault, Deleuze*. (Spatial Practices 7). Amsterdam, New York: Rodopi. https://doi.org/10.1163/9789042029132

Westbury, Chris 2005. Implicit sound symbolism in lexical access: Evidence from an interference task. In: *Brain and Language* 93(1), 10–19. https://doi.org/10.1016/j.bandl.2004.07.006

Whaler, James 1956. *Counterpoint and Symbol: An Inquiry into the Rhythm of Milton's Epic Style*. (Anglistica 6). Copenhagen: Rosenkilde and Bagger.

Wimsatt, William Kurtz 1954. *The Verbal Icon*: *Studies in the Meaning of Poetry*. Lexington, KY: University of Kentucky Press.