

Attention Tracking: Some Empiricism for a Visual Poetics

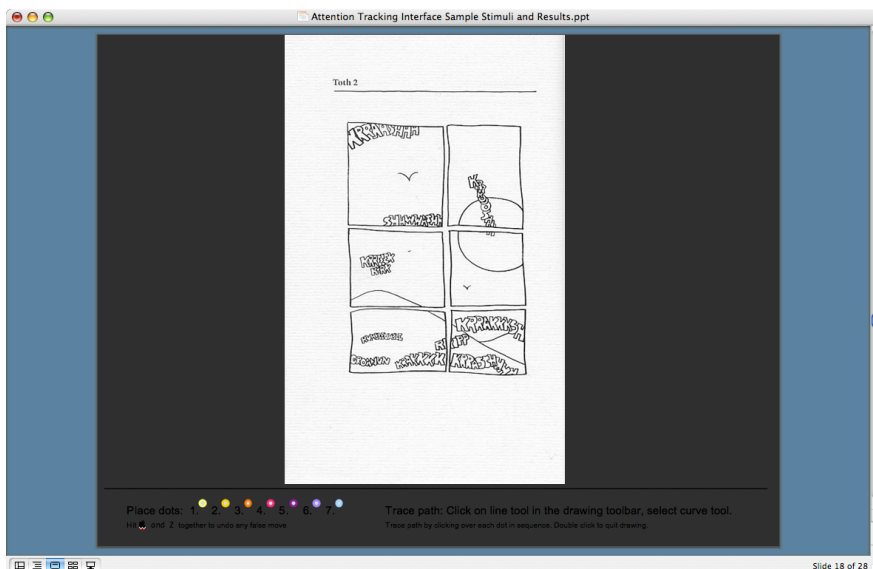
REVISION DRAFT DRAFT DRAFT DRAFT

(adapted from a chapter of *Visual Poetics: Meaning Space from Mallarmé to Metalheart*, Stanford PhD Dissertation, 2004)

ABSTRACT

In this paper I describe a simple empirical methodology for applying attention tracking to the literary study of visual texts. In particular I apply it to making sense of the notion “visual poetry”, as a tool to substantiate and explain “poetic” effects on the intimate cognitive grounds of readerly reception. The paper starts by identifying reading order and sequential processing as gateway questions for addressing visual language and visual poetry as concepts. It then takes up the more established methodology of eye tracking, reviewing the literature on reading and scene perception for insights relevant to an intermedial theory of visual reading. Describing also the limitations of eye tracking as a methodology, in particular the difficulty of concluding anything about mental attention and reading from the dry data on eye movements, I describe my attention tracking methodology as a radically simplified alternative, and argue for its superiority on this key point. I then report on an extended study conducted with Dr. Barbara Tversky of Stanford University, and demonstrate how the analysis of attention can be used to reveal textual structures, ground interpretations, and explain literary effects in a visual reading experience. The theory of reading that emerges from this analysis is one that can easily embrace word and image in a coherent account, laying the foundation for a mature reception poetics of visual poetry and intermedial textuality in general.

0. Introduction



Attention Tracking Interface with a slide showing a page by bpnichol.

Making sense of the notion “visual poetry” requires some effort. To common sense the term generally seems either contradictory, because poetry is verbal by nature and any visual dimension can be at most supplemental, or very unrigorous, where “poetry” is applied to a non-verbal work simply to mark some

unspecifiable extra value, like the cliché “poetry in motion” for figure skating. The marginality of visual poetry as a field or as a practice, at least in the minds of critics, has a lot to do with the difficulty of finding much to say about it. The articulations of conventional literary criticism, poetics and prosodics lack traction on visual texts precisely to the extent that their textuality is visual, and thus disappears from the optic of a linguistics-based analysis. Without wishing to dull the challenge that visual poetry poses to literary practice precisely along these lines, I would like to open a perspective onto visual literary production that can demonstrate at least some of the textual richness visuality of itself can generate.

The question comes down to how, and how satisfyingly, anything visual can be read as poetry. Conventional criteria for calling a text poetry – its lyrical or metrical character, the density, compactness or layerdness of its expression, its transgression of common usages, its use of metaphor and other figures – seem hard to apply in the absence of a linguistic base. This, in addition to the natural myopias of academic specialization, helps explain why literary discussion of visual poetry has so often misrepresented the field by choosing to focus on forms and instances at the verbal-most end of the intermedial spectrum, because that’s where their analyses hold.¹ But to give up on visual poetry at the point where it exits the verbal is not only to miss much of the genre’s challenge to poetry as a practice and as an experience; it is also to miss the challenge it poses to our understanding of language. For, before anything can be read as visual poetry in a rich sense, we have to acknowledge the richness of how the visual can be read. We have to acknowledge that the visual can be, and inevitably is, read, and consequently that, far from leaving language behind when it exits the verbal, visual poetry instead brings it with, into the visual.

A theory of visual poetry (and poetic visual effects) requires a theory of visual reading, because poetry is experienced not as recognition of an object, but in the experience of certain effects of reading. And a theory of visual reading in turn requires a theory of visual language, or an account of the visual meaning resources underwriting a text, one that will amount to a theory of language in a supra-ordinate sense, capable of substantiating the idea of reading intermedially. Key to the significance of visual poetic practices historically, and to the value of theorizing visual reading with their help, is that these call for visual language in practice, and as a consequence call it forth in the cultural blend of intermedial experimentalism. But how are we to engage this language and these languages critically, what analyses are we able to apply to the kinds of text they make? There is no end to the angles and lenses that can be applied to “reading” visual texts in art-critical, literary-theoretical or cultural-critical discussion, but the particular challenge of taking the “reading” of visual poetry literally, of thinking visual language minutely and observing it in behaviors, remains largely unmet. We have no

¹ The corpus of works for which claims to visual poetry are made covers a broad spectrum, from word-only texts with some degree of visual inflection (starting from, say, William Carlos Williams and e.e.cummings, and increasing through George Herbert, Eugen Gomringer and F.T.Marineti), through texts that mix verbal and visual elements (Klaus Peter Dencker, b.p.nichol and Julien Blaine are good examples), to works consisting exclusively of visual elements (e.g. Max Ernst’s collage poems, the “other” futurist Giuseppe Steiner or the recent Flash poetry of Rainer Strasser). A prominent example of this partiality in framing visual poetry is Willard Bohn’s *Modern Visual Poetry* (2001). Similarly, Marjorie Perloff received criticism from visual poets in the US for a 1997 graduate seminar on visual poetry at Stanford, in which she foregrounded slightly-visual work by predominantly verbal poets.

comprehensive theory of semantics to underwrite our taking visual form as language, so we have to look to the phenomenological immediacy of interpretive response for clues as to how possibly-poetic meaning is made of visual texts. Visual language, like visual reading, is not the province of a delimited faculty performing according to more or less known rules, but is rather the correlate of an activity that calls into play the full range of visual perception and intelligent response, including everything from lexical processing of verbal signs and iconic recognition of objects to the awareness of space and situations.

What I am presenting here is a tool for building a theory of visual reading in this broad sense, by enabling the observation and investigation of visual readerly experience. Attention tracking is an extreme simple method for investigating and reporting on how we read visual texts. In this paper I present it with an eye to making visible some of the dynamics that reveal what it can mean to say we read those texts as poetry. In this form it is specifically an articulatory tool for literary analysis serving a phenomenological visual poetics. But the generality of the reading that can be observed and analysed makes attention tracking a tool for applying these optics and analyses more widely, and before applying it as poetics, I argue its scientific basis, explain the empirical procedure, and discuss questions of analysis. Attention tracking is an awareness tool, both empirical and introspective, for tracking the reading experience, and pinning it, through the behaviors of attention, to particulars. It supplies any reader the starting point for a phenomenological reception study of visual texts. And however clinical it may sound, that is something very fundamental. Seeing as in a broad sense, for us, they say, 80% of the world comes in as visual, requiring reading.

The methodology does not in itself amount to a theory of visual reading, but rather serves to anchor any such to the phenomenological starting point, the moment of perceptual/cognitive encounter. From there I sketch only a few steps in the direction of a fuller theory of visual textual reception, but hopefully far enough to show how it can grow with the input of different discourses (poetics, psychology of perception/awareness, cognitive linguistics) into a highly articulated and rich way of describing the moment-by-moment construal of meaning from visual texts, also in terms that might convincingly account for the experience of poetic qualities in that construal. The methodology is empirically grounded, and so could potentially support a robust theory of visual literary experience, but it is also designed to strengthen first-person phenomenological observation, and so supports a visual literary and media education of the kind that could foster the development of an informed “readership”, critical reception and awareness.

I. Reading Order

Classically, reading and the viewing of pictures have been distinguished on the premise that one occurs sequentially and the other “simultaneously”. There are good reasons for arguing that more visual information

is taken in at a single moment of viewing a picture or a scene than at a single moment of reading lines of text², but aside from this it is misleading, and for the most part meaningless, to describe picture viewing as simultaneous. Picture viewing is every bit as sequential (or temporal) as verbal reading, if not as geometrically linear, in the sense that even a small image can only be satisfactorily perceived by means of multiple fixations of the gaze, constituting a total image only through a series of discrete perceptual acts. What may be experienced as simultaneous is the outcome of this process, a stabilized representation of the viewed object or field, which is however a mental construct, an internalized correlate of the physical object, and not the actual live input of perception. This internal representation, furthermore, is not necessarily more simultaneous or static than the equivalent representation constructed internally during verbal reading. A verbal text describing a static image, however sequentially delivered, will be constructed as a static, or simultaneous image in the mind, though even there, it would seem, a kind of internal scanning is required to perceive the various parts. On the other hand a painting, even while simultaneously present as a physical object to the view, will often be perceived as a pattern of movements, and so represented internally as such.

As Gyorgy Kepes explains it, the visual coherence of a picture is itself the product of sequential dynamics:

The final task of plastic organization is, then, the creation of an optical structure of movement that will dictate the direction and progression of plastic relationships until the experience reaches full integration...The kinetic basis of plastic organization—the linear paths of the eye on the picture-plane—is the common measure that binds into a unity the changing plastic relationships (Kepes 59).

The possibility of theorizing visual reading depends on the possibility of reconciling the procedures of sequential visual attention with the internal dynamics of information processing and “construal”. The interaction of these two processes defines reading, and because picture viewing is also sequential, the notion “visual reading” has none of the problematic quality present in the terms “visual language” or “visual poetry”. Even Nelson Goodman, whose firm distinctions between pictures and paragraphs try to exclude this hybridizing, has no problem saying that pictures are to be read (Goodman ???). The challenge of a visual reading theory that would answer the questions of visual poetry, however, is to observe in the reading process itself the crucial continuity between the verbal experience called poetry, and the visual one for which claims to poetry are made.

Many of the literary readings of visual texts I offered in my dissertation, *Visual Poetics: Spatial Meaning from Mallarmé to Metalheart* (Stanford 2004) to substantiate the idea of visual poetry in non-verbal materials, whether of David Arnold’s photo poems or of the 3D graphics texts of Metalheart artists Tim Jester and Per Gustafson, relied heavily on the role of visual features in determining a temporal order to at least portions of

² For example, the fact that for an individual focusing of visual attention, peripheral vision is effective in gathering meaningful information at a significantly wider radius when viewing scenes than when reading text; cf. Rayner, p.399.

the viewing. In Jester's "Blood", I observed a counter-clockwise flow that helps channel the eyes' movements in a pattern sharpening the semantic focus of the piece. And in both Gustafson's "Tribute to Metalheart" and Arnold's "and all the whining sounds" I pointed to effects established by delaying presentation of certain details until certain others have been perceived first, and the construal process carried far enough to effect a surprise or reversal on that basis. The possibility of such effects is vital to a robust conception of visual poetry as I want to portray it here, even if such effects themselves are not necessarily characteristic of the genre as it exists. They are important because of their potential to fulfill the requirement of a sufficient "language-likeness" in visual texts, as one way of substantiating the notion that pictorial signs can be organized to perform discursively, and within that discursivity do poetry.

Reading order is critical to the functioning of verbal syntax and to the directional build-up of meaning in verbal poems. So, one possibility for recognizing visual reading effects as poetry, is if visual texts can establish some analogous mode of ordering. Rosalind Krauss sees even a non-directed sequentiality as approaching Rauschenberg's "flat-bed" collages to the condition of language, generating "an undeniable experience of syntax" that shares "with language some of its character of discourse" (Krauss 40). Adding to this a more determinate temporal/spatial ordering should go even further, and may answer adequately to the intermedial artist Dick Higgins' condition that in a poetic work where the visual element predominates, "something of a verbal method of experiencing the work remains—the process of reading, of abstracting a sort of verbal pattern from the work ..." (Higgins 47).

Willard Bohn, a prominent scholar on the topic, sees reading order as a crucial aspect of effective visual poetry, though the judgement is self-selecting in his case, since the corpus he constructs in his *Modern Visual Poetry* is made up almost entirely of verbal-intensive examples. He argues at considerable length that the individual calligrammes that compose Apollinaire's "Paysage" were designed to be read in a certain order, and that that order both facilitates fluid reading of the parts, and purposefully constrains how the piece gets interpreted. Elsewhere, he criticizes poets for falling short of this basic requirement of visual poetics by failing to establish a clear reading order in their poems. While I think it limits our reading of "Paysage" to insist that there is a single right order in which to sequence the pieces, in the case of a predominantly verbal form of visual poetry looking to reconstitute a strict linearity in the lines of text may indeed make sense. But for the majority of visual poems, and particularly for the "very visual" range that concerns me here, it would be wrong to either expect or require anything so determinate. In winning for poetry a measure of independence from verbal language, the rules of linear sequencing are among the first to go.

Poets often insist on the value of visual poetry as a "non-linear" mode, and a chief virtue of being non-linear is allowing for a variety of reading orders. Virginia La Charité, in her book *The Dynamics of Space*, makes this point about *Un coup de dés*, though the spatialized layout there only minimally impedes following the poem as an obvious sequence of lines. The forms of visual poetry that emerge after Mallarmé include many in which any possibility of deciding a sequence for reading is studiously eliminated. The enthusiasm surrounding early

literary experiments with hypertext centered largely on the possibility of unlimited and unconstrained possibilities of sequencing. Much, even most, hypertext poetry, however, failed to yield interesting results precisely because the possibilities were left too wide open, making uninteresting compositions as or more likely than interesting ones. When Adrian Pilkington points out that successful poetic devices depend on establishing a “directed” search of weak implicatures, he is merely reformulating a basic lesson of 20th Century experiments with indeterminacy in art and poetry. The so-called LANGUAGE poets, for example, who sought to maximally challenge the ordering regimes of conventional language use and thus continually advocated indeterminacy as a poetic strategy, show themselves to be equally preoccupied with closure as a crucial ingredient. Ron Silliman concludes his essay “Migratory Meaning” with the observation that the primary need in formulating a contemporary poetics focused on indeterminacy, is an understanding of the devices that motivate closure. And Lyn Hejinian’s now classic essay “Rejection of Closure”, similarly argues that the trick lies in walking a fine line between total absence of constraints, and fatal fixity of meaning. She writes:

I want to say this at the outset and most emphatically, in order to prevent any misunderstanding. Indeed, the conjunction of *form* with radical *openness* may provide a version of the “paradise” for which the poem yearns—a flowering focus on a distinct infinity. (Hejinian 27)

2. Eyetracking

Speculating on the reading order of a visual poem, or on general principles of how visual texts are scanned and processed, is of little use without some means of verifying how viewers in fact view texts. Naive theories that assumed pictures are scanned left to right like text, or on a diagonal from upper left to lower right, or clock-wise by right-handed and counter-clockwise by left handed viewers (Saint-Martin, 189), have not been borne out by empirical studies. Eye-tracking technologies that allow precise recording of the rapid-fire movements the eyes make when viewing anything reveal no “normal” pattern of viewing. Eye movement research is slightly older than the modern tradition in visual poetry, dating back to psychological studies of reading from the latter 19th and early 20th Centuries (cf. Rayner 372). Since the 1970’s both the technology and methodologies for studying eye movements have developed dramatically, and extensive, sustained research has been conducted, yet still no simple or standard viewing procedure has been identified. Researchers today are still uncertain whether eye placement is decided more on the basis of semantic or perceptual properties of the candidate objects (Rayner 398-9).

Nevertheless, eyetracking would seem a promising means for confirming hypotheses as to how a visual text gets read. By recording the eye movements of a group of readers viewing Apollinaire’s “Paysage”, for example, we could statistically confirm, if not the order Apollinaire intended us to read its various short

elements in, at least whether there happens to be clear consensus among viewers in following a particular order. Of course, for such a simple example we could presumably rely on the readers' own memory of how they read the poem, and save ourselves the effort and expense of using the complicated devices and procedures of eyetracking. For more complex works, or where the question is not In which sequence were these four calligrammes read?, but rather Where did the eyes go in reading this piece, more sophisticated methods are required. The speed and number of eye movements easily exceed our ability to keep track of them ourselves. To confirm the optical flow I observed in Tim Jester's "Blood", for example, or to determine what visual elements were viewed and in what order, we would need precise data on where each subsequent fixation landed, and since eye movements as revealed by eyetracking technologies are notoriously jerky and stuttering, we would also need some means of statistically averaging the results to discount distracting data.

Though I am unaware of any previous study to apply the methodology of eyetracking to visual poetry, its possible relevance to graphic design has long been recognized. Despite Rudolf Arnheim's pessimism concerning the usefulness of such an approach to formalist study – "there are very few links between the order and the direction of the fixations and the compositional structure of a work" (Arnheim 364; qtd in Saint-Martin 189) – eye movement research has seemed promising for understanding the efficacy of information delivery in graphic design. In the 1980's the journal *Visible Language* devoted a two-volume issue to eye-tracking research³, most of it clinical work that still required extrapolation to possible applications for design theory and practice. Since at least that time, eyetracking has been a tool in the high-stakes business of attention-getting and messaging, serving to confirm or correct the visual strategy of glossy magazine covers and advertisements. More recently, the new field of webdesign, whose interactive variables are beyond the scope of conventional industry wisdom, has sought to benefit from the scientific observations of eyetracking. The California firm, Eyetools, Inc., for example, offers eyetracking services to companies eager to maximize the impact and efficiency of their on-line presence. They use the techniques of eye movement research to "capture and depict customer viewing patterns and behavior", and answer the vital client questions "Is advertising being viewed? Where am I losing customers? Do viewers recognize branding elements?"⁴

The two main areas addressed in eye movement research are reading and scene perception, offering insights relevant to both aspects of the intermedial "language" of visual poetry, though until recently, relatively little research had been done directly considering mixed verbal and visual forms of reading, e.g. cartoons or mechanical diagrams (Rayner 392). Recently the interest in user interaction with websites has spurred a great deal of such research. The basic picture we get from this research, applicable to both reading and viewing, is of a gaze in constant, irregular movement. The viewing gaze is literally in constant movement, even during the relative stillness of a fixation, due to a phenomenon called *nystagmus* which involves a continual

³ I'd like to thank David Arnold for a generous gift of books that included these two volumes.

⁴ From the Eyetools, Inc. website, <http://www.eyetools.com/new/applications/htm>. Accessed July 27 2004.

slight oscillation of the eye to keep renewing input on the retina (Rayner 373). The more significant eye movements, made in search of new information, are called *saccades* and occur on average 3 to 5 times per second (Henderson 260), interrupted by pauses for processing, called *fixations*, lasting an average of 200 to 300 microseconds each (Rayner 372). The field of vision, continually shifted and repositioned by eye movements, is divided into three concentric regions of variable acuity. The *focal* region, spanning about 2° at the center of vision (about the size of a thumbnail at arm's length), is the highest in acuity, surrounded by a *macular* or *parafoveal* region of significantly reduced acuity, extending approximately 5° beyond the focal core in all directions, in turn surrounded by the *peripheral* region of very limited acuity extending to the edge of vision, some 120° in diameter (Rayner 374).

The behavior of the eyes in reading shows many irregularities. Studies show that 10 to 15% of fixations actually go backward along the string of text, 15% of words get fixated more than once (Rayner 387), and about 40% of words in a text never get fixated at all (Underwood 118). These irregularities are due in part to the fact that looking and seeing, fixating and reading are not the same thing, and the key challenge in using eye movement studies to build a theory of reading is to discern within the blur of physical evidence about looking, something of the invisible processes of seeing and making sense.

One crucial distinction that emerges from the research is between where we are looking and where we are directing our attention. William James, and no doubt others before him, had already noted that we can direct attention to objects in the periphery of our visual field without focusing the eyes there (cf. Henderson 260). This manifests, in empirical studies of reading, in a distinction between the size and shape of the ocular focus (based on acuity factors in the retina), and what is called the *perceptual span*, or the “region from which useful information is acquired during an eye fixation” (ibid), based on the freedom of movement of attention beyond the foveal region. Studies since the 70's have shown conclusively that the perceptual span is asymmetric, stretching in the direction of movement (e.g. to the right when reading English) and beyond the scope of ocular focus. When the eye lands on a particular point in a word or sentence, it can discern up to three or four characters to the left of the one fixated, but up to 15 characters to the right (Henderson 261; Rayner 380). The *word identification span*, i.e. the span within which words can not only be discerned but also read, is slightly narrower, extending perhaps 7 or 8 spaces to the right (Rayner 380). These facts help explain how we might skip 40% of the words in a text and still understand it. They also explain how we might know ahead of time what words we can afford to skip; according to one study, we skip only 18% of “content” words but fully 62% of “function” words (Underwood 118). At the same time, most studies conclude that no information is gathered from below the line being read, suggesting that the reading span is purposefully shaped to the task and material conventions of reading. In reading, the viewer makes efficient use of this shaping, restricting attentive activity to the narrow slot of “incoming” text, and keeping attention slightly ahead of the eyes to help strategize effective placement of the next fixation. And when the eyes land on a word, they do so

in accordance with the asymmetry of the perceptual span, its momentum pushing on in the direction of reading: the *preferred viewing position*, or the position at which a word is most likely to be fixated, is about half-way between the beginning and middle of the word, only slightly off from what has been identified as the *optimal viewing position* given the forward-looking structure of attention (Rayner 385).

Acknowledging the divergence between eye movements and internal cognitive processing, a number of theories have been advanced to explain how phenomena such as fixation placement and duration relate to the task of reading (Rayner 388). Though countervailing theories argue that eye movements reveal little to nothing about internal processing, a general model of eye movement control emerges that provides a starting-point for an empirical theory of reading able to address both verbal reading and the reading of visual texts or scenes. Building on a framework laid by R.E. Morrison⁵ and further developed by several others, John Henderson elaborates what he calls a “Sequential Attention Model”. The basic assumptions of this model are as follows:

The sequential attention model contains five basic assumptions.... First, at the beginning of each new eye fixation visual attention is allocated to the stimulus at the center of fixation. In reading, the attended stimulus is likely to be the word..., though in the case of longer words it may be just one part of the word. In scene perception, it would presumably be at the level of the object. Second, attention is reallocated to a new stimulus when the foveal stimulus is “understood.” The simplest interpretation of “understood” here is that attention is reallocated when the foveal stimulus is identified.... However, attention could be reallocated when activation from the foveal stimulus reaches a critical threshold prior to recognition, or alternatively could be reallocated when a process following identification such as syntactic parsing (in reading) or semantic interpretation (in scene perception) is imminent or completed. Third, the reallocation of attention is coincident with two aspects of eye movement programming: (a) when attention is reallocated, the system begins to program the motor movements necessary to bring the eyes to a new location, and (b) the new locus of attention is taken to be the location toward which the eyes should be moved. Fourth, the reallocation of attention to a new location gates higher level analysis at that new location. Finally, the eyes follow the shift of attention to the attended location following the eye movement programming latency. (Henderson 263-4).

To summarize: Attention aligns with the foveal focal-point of the eyes to get high-resolution information from a particular spot or object in the visual field. When the information from that spot has been processed sufficiently (whether simply to the point of word/object-identification, or to the point of threading that word or object (as sign) into a larger project of construal) attention moves off of the spot even while the eyes remain there. Using something like a “preattentive map” (Henderson 264), a peripheral awareness of the rest of the visual field, attention chooses a new spot to focus on, and begins perceiving what it can there without the extra acuity of ocular fixation. At the same time, the oculomotor system receives neural commands to

⁵ R.E. Morrison, “Manipulation of Stimulus-Onset Display in Reading: Evidence for Parallel Programming of Saccades”, *Journal of Experimental Psychology: Human Perception and Performance*, 10, 1984, 667-682.

relocate to the position chosen by attention, and when the system is ready the eyes shift the high-acuity core of vision to align with attention and complete the information gathering from that spot.

Such an account, as Henderson and others make explicit (Henderson 266; Rayner 402; Underwood 116), rough as it is, should apply equally to reading and to scene perception or visual information processing more generally. In both cases, movements of the eyes and attention are taken to serve functions of sequential uptake, parsing and construal, and to offer a view-from-without onto the internal processing taking place at each stage. As such the model offers us a promising basis for an intermedial theory of reading. From the cognitive standpoint of eye movement research, reading and viewing are basically equivalent, or more to the point, viewing *is* visual reading. This equivalence would presumably be made even more apparent if, in addition to the bulk of work done on object recognition and “scene perception”, more research were focused on more semiotically “articulated” types of visual text, for example the pictogram writing I discussed in the first section. The differences between reading and viewing, as revealed by eye movement research, concern mainly the unique requirements of linguistic processing as opposed to “simple” perception and recognition, and it would be interesting to observe whether eye movements on articulate, “text-like” stimuli show up as notably more like verbal reading than eye movements on dense, “picture-like” scenes.

The possibility of using eye movement research to ground a theory of reading rests on some version of what is called the *eye-mind assumption*, which applies to both reading and scene-viewing, namely “the assumption that the direction of our eyes indicates the contents of our mind” (Underwood 111). The limitations of eye movement research for this purpose arise with the margin of error in that assumption, namely the discrepancy between the easily-perceived direction of gaze and the harder-to-perceive direction of attention. If attention can move independently of the eyes, and it seems it can wander even farther in picture viewing than in reading text, and not just left-to-right, but in any direction (Rayner 399), then the most meticulous recording of ocular fixations and saccades may or may not tell us anything about what gets read. As it is, even where the eye-mind assumption appears justified, the only direct link of any strength that can relate eye movement data to inner experience is the duration of fixation, which in normal reading situations is likely to respond to difficulty, or simply amount of processing (Underwood 111); hard words or words in hard sentences require longer processing times, and should correspondingly produce longer fixations, though extended fixation on a given word may also be motivated by retrospective processing of previous information, or even by integration of information noticed peripherally, but not yet fixated (113). Beyond that, eyetracking can tell us little about reading, without some means of observing the conceptual activity that accompanied the fixations; and what could even this tell us about poetry?

We can imagine an experiment observing fixation duration as an empirical measure of an elusive reader-response phenomenon such as “delayed categorization” or “extended exploration of weak implicature”, which Reuven Tsur and Adrian Pilkington, respectively, identify as central to poetry as a kind of

thought experience. Such an experiment might give us a behavioral angle on poetry as a cognitive effect or experience. But such an experiment would require an elaborate methodology of compensatory testing, probably relying on subject interviews or other retrospective reporting, to gain any insight into the crucial semantic aspects of processing, and because pictures tend to be informationally denser than text it would be exponentially harder to constrain for useful results in studying viewing as opposed to verbal reading.

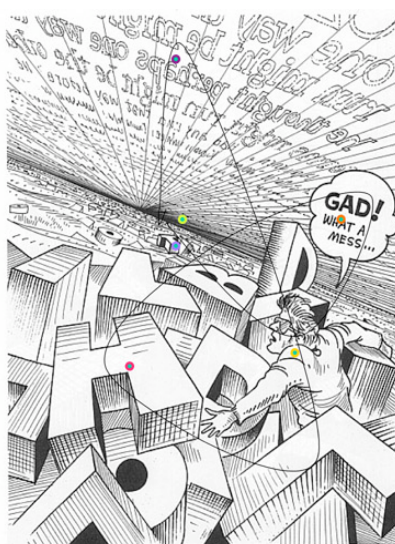
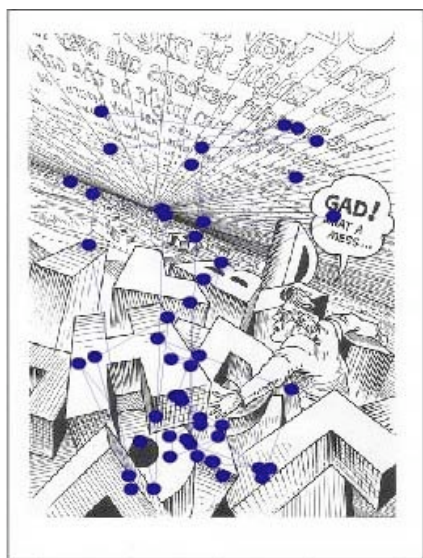
A number of researchers have questioned the usefulness of eye movement research to yield insights into viewing and scene perception at all. The fact that viewers, unlike readers of text, seem to “get the gist of a scene very early in the process of looking” has seemed to leave researchers with little to study, and to make eyetracking methodologies “a high-cost, low-yield endeavor” (Rayner 398). Furthermore the main liability of eyetracking (the marginal non-coincidence of eyes and attention) is presumably exacerbated in general viewing. In reading we can at least presume that attention to a word is attention to its meaning, whereas in viewing attention to an object or area could indicate attention to any range of perceptual properties, category identifications or semantic roles. If the “scene” or “picture” is a visual poem, however, the case is an intermediate one, where every object is to be taken as a sign, and we can never tell, especially after the very first fixations, whether the “text” was being viewed as a dimensional, scenic whole, examined as a flat field of objects, or read as articulated code for thoughts.

Even when verbal script is integrated into a scenic environment, as in the case of Arnold’s *Situations*, it has something of a different mode of appearing than objects or scenes as a whole. Isidore Isou, founder of Lettrisme, commented on this difference, observing that linguistic signs always overpower images and scenes. This may only have meant that words attracted his attention more than visual perceptual detail, yet there is a substantial difference in the “feel” of each class of signifiers. I observed that in Arnold’s work as a flicker between modes of perception. Through millennia of reading text, the visual experience has been one of dark marks on a white ground, excluding other visual variables from the area of focus in reading. The visual background, all depth dimension, has been whited out, so to speak, eliminated to aesthetically support the ideality of print transmission. The visual background noise, so to speak, of reading has been reduced to a strictly limited palette and texture of perceptual variables, discrete dark markings on a white-to-tan ground, or occasionally the reverse. Something of this experience remains in the characteristic feel of attending to textual components of a visual poem or scene, which almost inevitably elicit a sudden “nearing” or “focusing in” on the spot of the reading, as well as a different ratio of visual perceptual to inner conceptual attention in the experience.

Though a record of eye movements is clearly useful in determining viewing order, to observe anything about reading requires us to take further steps, confirming by performance in an experimental task, or by verbal report, whether and how a fixated area also received attention, and suggesting how information from that area might have figured in construal. As a result, eye-movement research must tackle very basic questions about reading if it is to have any strong results. According to the research review I have been citing

by Rayner, for example, through the late 90's a major question in eyetracking studies of reading was still what factors, informational or perceptual, had the greatest influence in guiding eye movements. Certain influential studies suggested that the most “informative” regions drew the eyes first, and others that out-of-place or incongruent elements did, amounting to much the same thing. Other studies, however, questioned these results on the basis that “informative” was often unintentionally confused in the studies with perceptually “distinctive”, and suggested that visual factors of distinctiveness and salience may actually be more influential (Rayner 398-9). Such questions and their answers are of course far from producing anything like a theory of reading at the cognitive level.

3. Attention Tracking



Eyetracking (left) and attention tracking (right) results on a page from John Riddel's *Criss Cross*.

The study I conducted with Dr. Barbara Tversky in 2001/02, on a grant from the Center for the Study of Language and Information (CSLI),⁶ a research institute affiliated with Stanford University, took a major open question of eye movement research as its point of departure. Interested in reading order and reading paths in visual poetry, I assumed that visual factors could play a determinate role in attracting and guiding attention within a visual text, and thereby in influencing the process of construal and the outcome of visual reading. By observing the viewing behavior of subjects “reading” visual poems, I hoped to discern patterns explicable by

⁶ This study, entitled “Spatial Meaning Constraints in Visual Reading”, was funded by a grant from CSLI in the first year of what is now the Media X project, focused on interdisciplinary research involving new media and interface technologies; see www.stanford.edu/mediax. It was conducted from September 2001 to September 2002.

the spatially constraining influence of visual factors. Where a majority of reader viewing paths overlapped, taking in the same visual elements in the same sequence, the piece would be revealed to have a *de facto* reading order, a layer of textual structure not visible to the “naked eye”. Furthermore the visual features that showed up as the most influential in constraining viewing order would by that token reveal potential *devices* for creating visual rhetorical or poetic effects.

In 2001 I had conducted a pilot study with similar aims. For that study I used a conventional eyetracking approach, showing 25 subjects a series of 17 visual poems on a computer display, and recording the eye movements and fixations they made as they read/viewed each piece. The results, predictably, were hard to interpret. The jerky scatter of fixation-points showed little strict repetition between viewings, and the majority of fixations fell ambiguously on spots not easily identifiable with distinguishable semantic units, nor immediately explainable on principles of visual salience. As in the example above, on some stimuli there was a marked preference for fixation at or near the edges of objects, but there seemed no consistent way of explaining what edges or what objects. Subjects reliably fixated certain types of elements, notably text, and in some cases there was substantial consistency in what objects of a particular piece got fixated even where the order was different. Elsewhere, particular regions of the field seemed to reliably guide viewing along roughly invariant paths, even where the sequence of fixations along that path was logged by different viewers at different points in the viewing process, contained different numbers of fixations, and did not touch on precisely the same “landmarks” along the path. It was these cases of sporadic and approximate invariance that seemed most promising among the results of the study, but without a more focused, statistical analysis to target the constraining power of, say, three-line convergences, maximum brightness areas, or other pre-selected features, the research questions could not be convincingly answered.

One reason eye movement research has not focused on visual poetry, not the main one, is that the difficulty of interpreting the results rises exponentially with the detail and complexity of the stimuli. To limit the complexity of the stimuli, or narrow the scope of enquiry, would have merely been to reproduce studies better conducted by psychologists. It would have been possible to narrow in on the perceptual constraints involved in determining first fixations, for example, but the issue of reading and of the semantic sequencing of visual elements would have had to be abandoned to make the eyetracking results more determinate. And there still would have been the problem of distinguishing between where the subjects were looking, and what they had in mind.

With questions as to the value of its results, eyetracking is a hard methodology to recommend for cross-disciplinary applications such as mine. The technology is extremely expensive and therefore hard to access, and requires sophisticated programming and statistical analysis to produce useable results. As an alternative to this, I devised a vanishingly simple method that would allow subjects themselves to record their reading path by marking up the on-screen image after an initial viewing. In the planning stage of our CSLI study, Barbara Tversky observed that this technique would likely give the results we were interested in, and

probably more clearly than in eyetracking data. Crucially, this technique, despite or partly because of the inevitable inaccuracies in subjects' reporting, promised to improve our experimental focus on what the readers attended, rather than merely what they looked at. Both for this reason and for the ease of collecting the data, we used this methodology for the entire experiment.

Though subjects would clearly be unable to precisely reproduce the sequence of their fixations, the error that would result would simplify in favor of those items that were actively attended. Instead of remembering exactly where they looked, we expected the results to show us what they took as significant, what they noticed consciously and involved in making sense of the piece, telling us more about the "text" as they experienced it than precise eye movement records could. In a number of eyetracking studies it has been shown that, while impossible to predict a subject's viewing order, the "scanpath" established on first viewing is encoded in the memory of a scene and tends to be repeated in subsequent viewing for recognition, even where subjects were not instructed to do so (Rayner 399). Thus there was reason to presume subjects would be reasonably successful in recalling the general path of their viewing. What would differ most from the data an eyetracking method would collect would likely be the number and specific placement of fixations. And since memory is classically a function of conscious awareness – we remember best what we were most conscious of – the likelihood of a subject remembering a particular fixation should correlate directly with the amount of attention and conscious processing they applied to that spot. What we would get, therefore, rather than a record of eye movements, would be a record of the movements of attention, the elusive goal of eyetracking methodologies. Attention tracking, as we called the procedure, would trade the technical complexity and objective certainty of eyetracking for a method any grade-school teacher can apply and a measure of subjective error that promises actually to reveal more directly than the precisions of eyetracking what most such research is ultimately after.

4. The Study

Stimuli

The stimuli I selected consisted of visual poetry and related visual art. The pieces, covering a range between very flat and very dimensional, were selected for their spatial articulation, or the degree to which spatial effects were integral to the meaning. Aside from Haroldo de Campos and Jiri Kolar, the visual poets represented were all "post-concrete", engaging visual space as such more actively than the poets of classical concretism or even much avant-garde work. Jiri Kolar's piece represented a transitional case between concrete and visual poetics and yielded interesting results. The Arakawa painting, while not generally associated with visual poetry, is a verbal/visual text thematizing spatial meanings and viewing patterns explicitly. And Duchamp's "La mariée mis a nu", even further from association with the genre, demonstrates attentional

guidance and discursive, textual structure in a visual art context. I included a total of six pieces from David Arnold's *Situations* to prepare a test case for studying the role of attentional guidance within the "very spatial" work of a single artist. And I added three slides representing architectural work by Arakawa and Gins, one a 3D computer rendering of an abstract space, and two photographs of built spaces. The Reversible Destiny project from which they derive activates spatial meanings via architecture as devices within a "built discourse" that represents the extreme development of spatial poetics emerging from experimental trends in 20th Century art and literature.

Procedure

The testing took place in two phases, an initial viewing phase and a repeat viewing and reporting phase. First, subjects were shown a Powerpoint slideshow containing each slide in the order they would appear in the reporting phase. They viewed each one for 10 seconds to familiarize themselves with the piece.

After previewing all of the slides, subjects were instructed in the use of the simple interface I designed for reporting their viewing paths. They were told to view the slides again, this time marking the path they followed using a curved-line tool in the Powerpoint tool palette, and then placing up to seven dots on spots they remembered looking at, in order, starting with the first. The dots were color-coded to mark the numerical order.

For the first eleven slides these were the only instructions. For the rest, there was an extra step of verbal reporting. For five of the David Arnold pieces, subjects responded briefly to two questions after each slide: "Please narrate where you looked from spot to spot?" and "What was most interesting about the space?" These questions were designed to elicit verbalized corollaries to their graphic reporting, allowing me to confirm what subjects were looking at for a given dot-placement, and to observe how features or objects were identified and preliminarily construed. For the abstract structural piece by Arakawa and Gins, and for the sixth David Arnold piece, the follow-up prompt was switched to "Describe the scene in your own words". And for the last two Arakawa and Gins slide, instead of tracing their viewing path on first revisiting the slide, subjects answered the question "If you were here, where would you go?" using the curved-line tool to answer, and in a follow-up slide responded verbally to the question, "Why?" The choice of questions for follow-up, verbal reporting can make a great difference in what kind of information is gleaned from the graphically reported attention paths, offering the possibility of comparing two very different modalities of response. In this study, the last variation proved unclear and did not produce useful results.

Data

The data produced took the form of graphical records of the subjects' recollected viewing paths, plus

the verbal reporting from the later slides. Each treated slide represents an individual “reading” of the piece, and collectively, a composite of all readings of a single slide provides a statistical indicator of the piece’s structure as experienced by a community of readers. By limiting the marking task to a maximum of seven dots, I necessarily constrained the results to reveal only the initial portion of the viewing. In considering the possible perceptual factors affecting attention, it is these early moments that promised to be the most revealing, as later exploration is presumably increasingly motivated by developing thought and observation patterns unrelated to gestalt perceptual cues. The same method, however, could be used to study an extended reading, though in that case it would be important to design the interface such that dots and lines disappear from the subject’s screen during marking, so as not to crowd the screen and distort viewing (a revision that would also improve the present study).

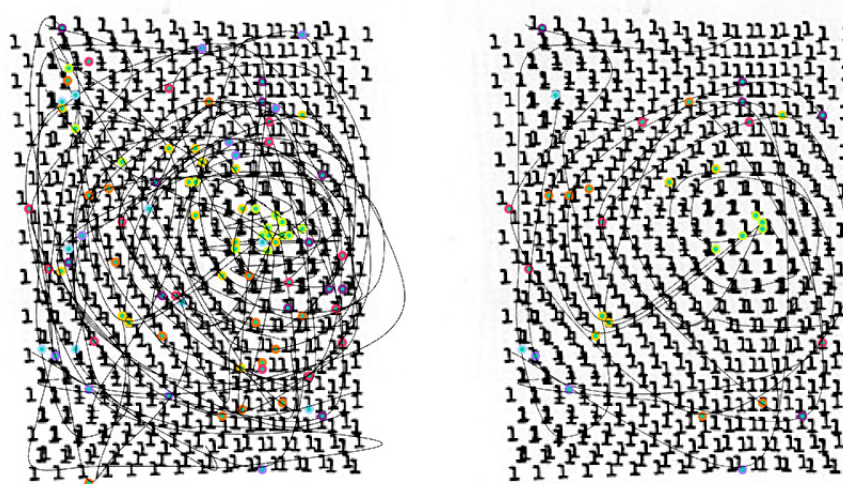
The first question of interest in reviewing the data is whether the viewing patterns appear random or motivated, and if motivated, by what factors. Jiri Kolar’s piece, “Homage [tribute?] to Ladislav Novak” presents a convenient case for studying the role of strictly perceptual factors, as the wavy field of i’s is what Steve McCaffery would call a “proto-semantic” construction; it presents nothing to read at the lexical level. The first image presented below is a composite of the viewing paths recorded by all 21 subjects. While chaotic, certain patterns stand out even to an initial visual analysis. Clearly a majority of all first placements coincide in the region, center right, where the letterforms are most widely dispersed, creating the impression of a raised, spherical form. A slightly narrower majority of second placements are concentrated in a region up and to the left at an angle of about 45° from that focal center, in a direction accounting for all first placements not falling in the central region. And most third placements landed directly below the region of majority first placements, though percentage deviation increased further between second and third placements. The other notable pattern is that most viewings took a clearly circular pattern, moving off that first region and circling around it; figure 2 below shows the counterclockwise viewings.

Clearly the subjects responded to this piece as a specific topography, a field of material resistance, and not as a screen of white noise. The three regions mentioned consistently drew early attention, regardless of the exact order. Though the pattern was more likely for viewers who scanned in a counterclockwise trajectory (2), the constraining influence at work at those points becomes even more interesting when we notice that some of the viewing paths with nearly identical results for the first three placements were moving in opposite directions (3,4). In the absence of distinctive objects at those locations, we have to consider the field properties of density and shaping. The piece creates the spatial impression of a field of raised spherical embossures and sunken troughs between them. What the viewing records suggest is that the gaze is engaged in construing this shaping *as terrain*, and thus follows paths responding to visual features as constraints on imagined movement.

A full statistical analysis of the viewpaths, based on an information analysis of the stimulus field, would be required to formalize the factors prompting attentional response, e.g. relative density, rectilinear versus

curvilinear alignment, formal salience over a given area and apparent depth or elevation. The central lesson for a strategy of visual poetic composition, simpler than that, lies in the connection between viewing and imagined moving, which I will return to at the end of this chapter. As regards this particular poem, however, the movement itself turns out to be a meaningful interpretive response, when we relate the piece to its literary reference. “Individualista” is the name of a concrete poem by Ladislav Novak. It consists of a rectangular, strictly rectilinear grid of i’s in the same typewriter font we see in Kolar’s poem. One of the i’s, at an unremarkable position in the grid, is turned upside down, in what the title leads us to read as a gesture of individualism and non-conformity. Kolar’s piece, by recasting the i’s in a more liquid formation redefines identity and social relation as phenomena of field dynamics. Each “individual’s” position is distinguished along smooth gradients of proximity and overlap, rather than according to the rigid assignment of coordinates. The field of individuality is “dense”, to extend Goodman’s terms, rather than “articulate”, and characters occupy all possible positions and degrees of relatedness with each other. The prominence of the embossed regions appears a property of fluctuations in the field, subject to fate and responsive to nature in a way refused by the militarized grid that provokes Novak’s character’s act of resistance. The viewing paths, by noticing certain regions over others, individuate the field, yet respond fluidly and non-hierarchically to its contours. The circular movements of attention the piece invariably provokes are themselves a refutation of the authoritarian coordinates of Novak’s grid, not to mention of the linearity of conventional text, and convey in the very process of reading at this proto-semantic level, an “exemplificative” reference to the subtle blendings between separateness and connection, individuality and crowd.

5. Analysis: Reading Reading Order

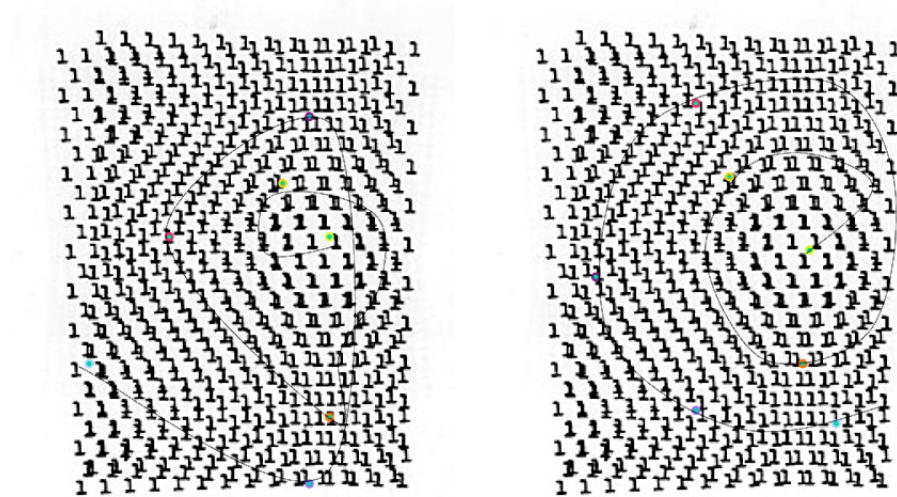


Composite of 21 viewings

Clockwise only



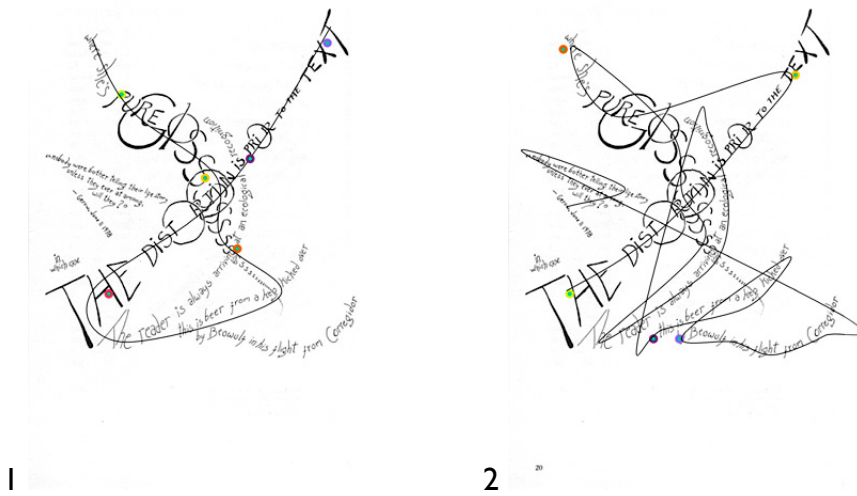
Where the visual elements are semantically invested, we get something more explicitly informative. As attention takes up unit after unit, we observe something equivalent in visual terms to the “chaîne parlée” or “verbal chain” (Saussure) in linguistics, the sequence of signs syntactically ordered and available for decoding. We might call it the “chaîne visée” or “attentional chain”, though as I will continue to explain shortly, the parallel is of course not exact.



In the case of a spatialized verbal poem, evidence of the order of reading has obvious value in gauging the determinacy of the verbal sequence in the composition. In reading the page from Jed Rasula’s *Tabula Rasula* (1986), below, just as many subjects started from the large text at bottom left as did from the text beginning in the upper left. Both approaches can be seen as accommodations of the top-down, left-right reading conventions for standard text, though other visual factors also played a role: a number of the subjects starting at upper left registered their first gaze in or adjacent to the large circular form of the “G” rather than at the actual start of that line, and nearly as many subjects looked first at the central conjunction of the main lines, where two adjacent O’s serve as a similar focal attraction, as did at either probable sentence beginning.

The significance of the indeterminacy in where to start reading comes down to the possible differences in meaning between the two ways of sequencing the lines. Would Rasula care whether his readers parse the page as “The distortion is prior to the text/ where she’s pure gloss”, or “Where she’s pure gloss/ the distortion is prior to the text?” More likely the ambiguity is evidence of a successful device employed by the author, mixing signals to remove reader certainty: “Where she’s” is highest and leftmost, but “The dist..” is actually further left and larger. If not, it is evidence of a failure to constrain attention more precisely. Either way, the ambiguous reading order is part of the distortion the text says is prior to itself, wrapping the reader in a knot that cannot be untangled from without. For this particular poem the simultaneity of possibilities is

more important than the question of where one is supposed to look first, but the same strategies are available for a case where precise order and timing are more crucial.



Where the significant elements are all or primarily visual, the importance of viewing order is not so clear. The more semantically articulated the visual elements, the more like reading we can expect the viewing to be. Certainly the pictograms in Lars Arrhenius’ “The Man without Qualities” are meant to be read “in order”, and it is interesting to note how readily we project the expectations of left-right, top-bottom reading into the visual context. Visual texts not so explicitly sequenced present more difficulties. Duchamp’s “La mariée mise a nu”, for example, casts the question of reading order in a different, and very interesting light when we consider the discursive semantics undergirding its visual forms. Though wordless itself, every element in the piece has a verbal label, representing a focus of ‘pataphysical references elaborated in the characteristically Duchampian musings of his “box” projects, notes and interviews⁷. This supporting structure of verbal meanings, no less suggestive or resistant to interpretive closure than the imagery, amplifies our experience of the work, particularly when charged with the non-random specificity of a subjective reading order. Translated according to Duchamp’s own labels, the perhaps confused viewing experiences of this enigmatic artwork yield articulate, if no less enigmatic texts:

1. Chocolate grinder necktie/ malic mould of the stationmaster/ crossing the horizon, the Bride’s garment/ (unlabeled support behind) the Bride/ central draught piston/ back through the Bride’s garment at the vanishing point of perspective/ scissors

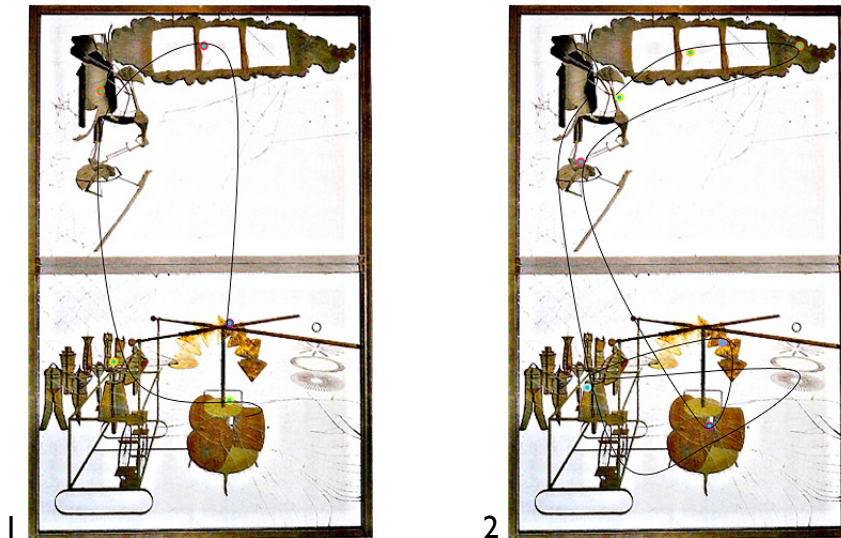
or

2. The Bride’s stem/ central draught piston/ milky way/ the Bride’s wasp/ crossing the horizon, the Bride’s garment/ chocolate grinder rollers/ sieves (or parasols)/ towards the malic moulds/ past the ocular witness, around the

⁷ The list of labels I use here was compiled by Jean Suquet, in his *Miroir de la Mariée* (Flammarion).

chocolate grinder/ up past the revolution of the bottle of Bénédictine in the sleigh/ malic mould of the waiter's assistant/ crossing the horizon, the Bride's garment again/ past the wasp, towards the Bride.

The second reading is clearly more spiritual.



6. Attentional Analysis

Barring a sufficient immersion in the Duchampian esoterica, such a demonstration may be more entertaining than informative. There is obviously a fallacy involved in assuming, outside of certain highly specialized cases, that visual elements as signs align in a definitive mental order once fixated, that a sequence of visual attending amounts to the firm syntax of words concatenated to form sentences. If the *chaîne parlée* is a one-dimensional sequence of words strung end to end, the *chaîne visée* is a two- or three-dimensional threading of distributed points into a fabric, looping, crossing, and linking in many directions. Though specifics of the temporal sequencing, and qualities of the movement of the gaze (direction, arc, rhythmical patterning) can play a meaningful role in the textual experience, the more basic question of syntax comes down to the network of relations established in a free exploration. Kepes and other commentators on compositional theory in the visual arts stress this non-linear, exploratory character of visual reading:

The ultimate aim of plastic organization is a structure of movement that dictates the direction and progression toward ever new spatial relationships until the experience achieves its fullest spatial saturation. As new relationships progressively unfold, the spatial integration of the image gains momentum until it finds final clarification in the plastic image as a whole. (Kepes 52)

Saint-Martin, observing that eye movements are of limited value in elaborating a visual semiotics from the perspective of the viewer/reader – “since these movements are potentially infinite and open to aleatory unfoldings leading to opposite conclusions” (Saint-Martin 188) – similarly turns her focus to the pattern of interrelationships established through free exploration:

Visual semiotics proposes that the composition, or rather the structure of the work, can be deduced only from a series of equilibria established between the elements. The [optical] energies and the regions they form, which are taken up and modified with a view to producing superior and more complex equilibria, can finally produce a state where the ensemble of movements and transformations produces a system which can then be offered provisionally as an adequate synthesis. (189-90)

While basically a restatement of Kepes' point, Saint-Martin articulates this notion of compositional structure beyond the basic gestaltian observations. For her, the “ocular circuits” manifested in viewing, rather than an additive concatenation of significant units, reveals a topological structure of relationships. The notion of topology allows Saint-Martin to consider the spatial structure of viewing without positing a definitive order or measure to that structure. While any and every spot in the visual field may receive attention, in any possible order, and while every perception associates the current spot of attention with everything previously attended, at least to test it for relations, perceptual equilibria (balances of light and dark, relative positioning, etc.) eventually establish a relative invariance in the structure experienced. Tensions and relations become apparent and emerge as the structure of the piece. Saint-Martin, true to the boundaries she draws for visual semiotics, discusses this process only at the level of primary perceptual processes, but clearly something similar happens at the semantic level, with items linking up in a network of possible iconic relationships and conceptual associations. Kepes makes this point explicitly:

As one searches for spatial order, and through the interrelationships of the plastic forces creates a unified spatial whole, one also searches for a meaning-order and builds from the different association-directions the common, meaningful whole. (Kepes 202)

So, except in special cases, what attention tracking reveals should not be thought of as the reading order, but rather as the attentional terrain of the piece in a viewer's experience. Since my experiment concentrated on the viewer's first few moves in each piece, the examples I have given reveal only the most salient aspects of this structure. Another approach would be to let viewers keep looking and marking their attentional path until they feel they have seen everything. Most likely this would not cover the screen with dots and lines uniformly; rather a distribution would emerge thoroughly marking the topology of interesting and uninteresting areas. The frequency of returns to a certain region would give a topographical articulation of this space, registering intensities of interest. Most likely such a map would reinforce the predominance of the first items attended, though it might equally reveal the importance of items only noticed later.

The first useful information from such an attentional analysis is simply whether a particular element was or was not attended. Where items consistently evade attention, they can hardly be said to figure in the reading of a piece, though unattended items can impact a reading peripherally or subliminally. Looking to the test results for a confirmation of my reading of David Arnold's pieces, for example, I discover that in the case of "this falling, tumbling, etc" (Prohm 2004 40-42) I was wrong, at least as regards the five white dots at the end of the hallway. While nearly every subject focused obediently on the bull's-eye of spray-painted rings, only six out of twenty-one subjects recorded a gaze on or near the spot in question. In the written follow-up, only two subjects (one of whom had not marked the spot) mentioned the dots directly, referring to "the three [sic] white dots in the background", or "the far dots of light". Two others mentioned the "hall" or "corridor", while one other identified "dark doors at the back". What spatial effect or meaning these few subjects perceived the dots to have is largely irrelevant to a formal reading of the piece, since statistically almost no one took them as significant elements. If everyone who did notice them had interpreted them the same way, we might argue that the detail is meaningful but hidden, and likely to be noticed by more subjects upon longer viewing, but the one subject I asked about them read them precisely the opposite way, i.e. as deepening rather than collapsing the hallway. Admittedly, there was no elaborate interpretation riding on the issue of these few dots, but the example goes to show how subject-specific perceptual responses can be, and how useful an attentional analysis that can take a larger community of responses into account



David Arnold, *Situations*: "this falling, tumbling, etc"



"because"

A second level of information revealed by this kind of method involves seeing whether one particular item was associated with another. If both items were attended, we know the association was possible, but as it turned out in the last case, we need the confirmation of a verbal report to be sure the association happened. In the case of the David Arnold piece, "because", where I read the central forest scene as an ambiguous element, temptingly interpretable as a real space, but possibly or probably only a poster, my reading was widely confirmed. Unsurprisingly, almost everyone (19 of 21) focused on the scene in question, which was at the center of the piece, though they did so at widely varied points in the temporal order. Of the 14 subjects who responded to the follow-up questions, 12 mentioned the scene among the "most interesting"

features. So, there is no question as to whether it figured in their reading of the piece. But my reading depended on two other factors: the parsing of the scene itself, and the syntactic relationship that set it in contrast or conflict with the surrounding environment.

Here the verbal reporting proves very useful, in that in addition to indicating whether a subject noticed the scene, the words used to refer to it reveal how it was identified. Of the 13 subjects who mentioned the scene in response to either or both of the follow-up questions, six clearly perceived it to be a poster, while five seemed to take it as a view to a real exterior, and two acknowledged it was ambiguous. Descriptions in the first case used the words “image”, “poster” and “cutout of the natural scene”. The responses on the “real” side varied in how definitively they indicated the subject’s interpretation. I list the expressions here in decreasing order of certainty:

1. “the hole in the wall” / “the row of trees and the bridge you can see through the wall”
2. “window through to the distance”
3. “window”
4. “the trees in the background”
5. “what appears to be a window”

The two unsure subjects each made a different first guess, one referring to “the picture – or is it a window” and the other asking “was it a window or merely a photo?”

So, the ambiguity I ascribed to the poster/window is statistically substantiated. But my reading of this piece relied also on a particular valuation of that ambiguity in the context of other elements. I argued that the question of whether it was a real place or a poster made a difference because of the claustrophobic decay around it, that this semantic framing drew a particular emotional investment in the scene and its possible identifications. I even made the claim that the decay made the viewer *want* to interpret it as real. One way of seeking confirmation for this would be to investigate whether subjects who paid more attention to the elements indicative of decay (e.g. the rubble on the floor) were more likely to read the scene as real, or precisely the contrary. The results, particularly from the attention tracking itself, are inconclusive on this, though of the six subjects who directly mentioned features of decay in their verbal report, pointing to “trash”, “messiness”, “graffiti”, “chunks of wall on the ground” etc, only one of them identified the scene as a window, while one saw both possibilities.

While the results do not support any statistical conclusions as to the prevalence of the reading I offered – a different approach with the follow-up questioning might have yielded more – verbal reports from two of the subjects strongly confirm my reading. One subject, narrating his experience of the piece, wrote:

The sylvan scene in the center of the frame trying to figure out how it fit, was it a window or merely a photo, the contrast with the tidy peaceful forest creating marked tension within me when contrasted with the disturbed and

disturbing space which frames it

The other, asked what about the scene was most interesting, answered:

The tension that arises from untainted (the picture of woods) clashing with the disorganized room

These responses at least show that the tension I identified as the semantic core of the piece is there as a possible content, and in the similarity of their phrasing suggest that the perception may have been more widespread than reported.

I did not include “and all the whining sounds”, David Arnold’s “Halloween” piece, in this study, and so cannot comment on the objective validity of my reading there. That reading relied on the timing of perception of the upside-down text along the baseboards of the room, the impulse to invert the picture to read the text, and the changed range of associations surrounding the rectangle of light on the floorboards once the picture was inverted. Reading behavior of this sort was beyond what the testing procedure would allow me to observe, but I want to make a few comments on the notion of time-specific reading effects before concluding this section. The first aspect of my reading would have been easy to test, simply by observing whether attention to the baseboard text was consistently logged after the other, more prominent features I enumerated. Compositional strategies for hiding or delaying detection of an item are obviously just the inverse of strategies to assure that an item is noticed right away. And, presumably, if it is crucial to a particular rhetorical or poetic effect that an item be perceived after one item but before another, within certain limits of precision the appropriate strategies can be arranged.

Meter and Rhythm

Although visual reading neither relies on nor limits itself to strict linear orders of the kind essential to verbal reading, temporal sequences and effects are both possible and pronounced in visual texts. One kind of significant temporal effect that does not depend on precise viewing order is visual rhythm. The page from Klaus Peter Dencker’s *Wortköpfe* below shows a distribution of elements that sews metrical qualities into any pattern of viewing. The space is asymmetrically filled, with a concentration of items in the lower, and particularly lower right portion. An approximate diagonal from lower left to upper right divides this region from the emptier upper portion, characterized by larger figures and more continuous lines. Figure 2 below shows the viewing paths of the first seven subjects, which, typical of the results as a whole, highlight the different visual weighting of these two regions.



1



2

With rare exceptions, everyone viewing this piece recorded their first gaze at one of two locations, either the large segmented circle at upper left, or the globe-head figure lower to the right. Of those who looked at the circle first, most looked at the head next, but no one who looked at the head looked next at the circle. Invariably two or three further fixations went to items in the lower right portion before attention again found its way up into the “quieter” region of the circle. This suggests two things. First, while viewings followed widely varied patterns, a basic current drew attention downward to the right and kept it occupied in that region, from which it might continue only later to explore the less dense region in the upper half. Secondly, it suggests that the head stands in “tighter” relationships than the circle, characterized by a greater number of near-neighboring attractors, and shorter distances of movement from one to the other, than is the case for the circle.

A gaze to the circle is a relatively leisurely gaze. Attention landing there finds less to draw it away, partly because of the focusing effect of the frontal, radial form of the target, partly because of the lines and arrows which do not draw attention so much as offer it traction, and partly because of the distance to cover between the circle and any next point of interest (other than the owl). Conversely focusing on the head, which does not meet the gaze but deflects it in the direction of the clustered elements, immediately leads to a series of fixations elsewhere, on the particular continents (labelled with names of computing languages) covering the scalp/globe, on the portions of text, or the other salient graphical elements. The effect of this difference, on a gaze that will eventually move back and forth between the two regions as it begins re-visiting items already focused on, is a metrical alternation, between close, quick movements, and a more singularly-directed, protracted gaze to the circle or empty region around it. The pull downward and to the right, even beyond the initial viewing, will be noticeable as a visual weighting whether the gaze at a particular moment is moving with it or against it, and the relative rest or relief of the upper region will counterpoint this, whenever it enters the stream of attentive acts. As the various visual elements are discerned and associated, the topology built up among them will contain these dynamics as part of its articulation. The scanning procedure

as it moves from spot to spot, in whatever particular order, will activate them as a meter or rhythm, establishing a baseline of alternating aesthetic qualities to underlie reception of the semantic units, just as sound patterning does in verbal poetry, though more random in its sequencing.

If a kind of meter in visual poetry is conceivable, a sequencing of beats iambic or trochaic, anapestic or dactylic according to the reader's own timing, a more obvious temporal effect is the general flow of viewing among elements. Attention tracking revealed the inherent circularity of viewing in Jiri Kolar's piece, and I suggested a possible thematic relevance for that. Similarly I pointed out a clear directionality in Dencker's piece, drawing attention directed at the upper region down toward the concentrated materials in the lower right. The attentional topology revealed by analysis is not neutral with regard to movement. It contains vectors and conductive potentials that structure the viewing and bias construal. If there is any thematic significance in patterns of attentional flow I observed in Dencker's piece, I would suggest it concerns iconic contrasts that parallel the formal ones. The globe-head, more fully rendered and dimensional, more detailed and "loaded", iconologically bearing the whole world and the bulk of information technologies, is impressively serene, yet more compressed and burdened than the broad, thin outline of a head containing the segmented circle. The one stands out as figure, while the other disperses as field. The one looks down, the other up, buoyed peripherally by arrows and a focus on the perched owl logo, suggesting wisdom rather than processing power. In this context, the downward draw on attention has an effect of gravity, and movements against that pull acquire the semantic valence of release or relief, coming into more spacious territory where the pull is upward and away.

Through analyses of this kind, we can begin to imagine a kind of literary criticism that would engage visual texts in visual terms, not merely as "enhanced" or "materialized" verbal texts, but on the basis of perceptual qualities and semantic functions operating in the visuals. Where visual perception supplies processes of complex semantic construal, seeing is reading, and the visual field *de facto* a site of language, whatever the materials. And identifying the process of construal, rather than any category or structure of signs, as the critical factor for a possible poetic experience, a theory of visual reception is well positioned to examine whether and to what extent "poetry" may be claimed of a purely or predominantly visual experience. The possibility of taking the term "visual poetry" both seriously and literally depends on evaluations of perceptual experience, which in turn depend on articulations supporting the detailed semantic analysis of such experience. Attention analysis and the visual reception theory it supports help bring us through visual perception to the same domain of conceptual processing where poetic experience of words takes place. It thus allows an integrated assessment of verbal and visual reading, in what I would argue is a more natural, less reductive approach to meaning. If anything, intermedial arts show us the ultimate coinvolvement of linguistic and non-linguistic codes, and a theory of meaning adequate to the demands of an intermedial poetics is also one that answers the challenge art in the intermedial context poses to compartmentalized semiotic discourses.

Acknowledgements. This research was funded in part by a grant from the Center for the Study of Language and Information, at Stanford University (2001-2). Editing and publication supported by a grant from the Finnish Cultural Fund (Suomen Kulttuurirahasto 2008).

Bibliography

- Apollinaire, Guillaume. *Calligrammes* (1918). Paris: Gallimard, 1970.
- Arnold, David. *Situations*. San Francisco: Trike Press, 1984.
- Barthes, Roland. *Elements of Semiology*. Trans. Annette Lavers and Colin Smith. New York: Hill and Wang, 1977.
- Berthoz, Alain. *The Brain's Sense of Movement*. Cambridge: Harvard University Press, 2002.
- Bohn, Willard. *Modern Visual Poetry*, (Newark: University of Delaware Press, 2001)
- Byrum, John and Crag Hill, eds. *CORE: A Symposium on Contemporary Visual Poetry*. Cleveland, Mill Valley: Generator Score, 1993.
- de Campos, Augusto. Questionnaire response. *Experimental/Visual/Concrete: Avant-Garde Poetry Since the 1960's*. Eds. K. David Jackson, Eric Vos and Johanna Drucker. Amsterdam – Atlanta: Rodopi B.V, 1996, p.376.
- Cohn, Robert Greer. *L'Oeuvre de Mallarmé, Un coup de des*. Paris: Librairie Les Lettres, 1950.
- Cubadust. *The Versus Project*. 2000. <www.cubadust.com/versus>. Accessed June 30, 2004.
- Drucker, Johanna. *The Century of Artists' Books*. New York: Granary Books, 1995.
- . "Experimental/Visual/Concrete", in *Figuring the Word*. New York: Granary Books, 1986.
- . *The Visible Word: Experimental Typography and Modern Art, 1909-1923*. Chicago: University of Chicago Press, 1994.
- Garnier, Pierre. *Spatialisme et poésie concrète*. Paris: Gallimard, 1968.
- Glazier, Loss Pequeño. *Digital Poetics: The Making of E-Poetries*. Tuscaloosa: University of Alabama Press, 2002.
- Gomringer, Eugen, ed. *Konkrete Poesie*. Stuttgart: Reclam, 1972.
- , ed. *Visuelle Poesie*. Stuttgart: Reclam, 1996.
- Goodman, Nelson. *Languages of Art*. Indianapolis: Bobbs-Merrill, 1968.
- Gustafson, Per. Interview with Jens Karlsson. *Nervous Room*. Ed. Jens Karlsson. <<http://nervousroom.com/interrogation/019/main.php>>. Accessed June 30, 2004.
- Helfand, Jessica. *Screen: Essays on Graphic Design, New Media, and Visual Culture*. Princeton: Princeton Architectural Press, 2001.
- Hejninian, Lyn. "The Rejection of Closure", in *Onward: Contemporary Poetry and Poetics*, ed. Peter Baker. New York: Peter Lang Publishing, 1996.
- Henderson, John. "Visual Attention and Eye Movement Control During Reading and Picture Viewing". *Eye Movements and Visual Cognition: Scene Perception and Reading*. Ed. Keith Rayner. New York: Springer-Verlag, 1992. 260-283.
- Higgins, Dick. "The Strategy of Visual Poetry: Three Aspects." *Visual Literary Criticism*. Ed. Richard Kostelanetz. Carbondale: University of Southern Illinois Press, 1979. 41-50.
- Iser, Wolfgang. *The Act of Reading: a Theory of Aesthetic Response*. Baltimore: Johns Hopkins University Press, 1978.
- Jakobson, Roman. "What is Poetry?" in *The Semiotics of Art*. Eds. Ladislav Matejka and Irwin R. Titunik. Cambridge: MIT Press, 1976, pp. 164-75.
- . "Closing Statement: Linguistics and Poetics". 1960. *Style in Language*. Ed. T.A.

- Sebeok. Cambridge: MIT Press, 1960. 350-377.
- Kepes, Gyorgy. *Language of Vision*. Chicago: Paul Theobald, 1941.
- Krauss, Rosalind. "Rauschenberg and the Materialized Image". *Robert Rauschenberg*. Ed. Brandon W. Joseph. Cambridge: MIT Press, 2002.
- La Charité, Virginia A. *The Dynamics of Space: Mallarmé's Un coup de dés Jamais N'Abolira le Hasard*. Lexington: French Forum Publishers, 1987.
- Lupton, Ellen. "Visual Dictionary". *The ABC's of : The Bauhaus and Design Theory*. Eds. Ellen Lupton and J. Abbot Miller. London: Thames & Hudson, 1993.
- Mallarmé, Stéphane. *The Collected Poems*. Trans. Henry Weinfield. Berkeley: University of California Press, 19??
- Mallarmé, Stéphane. *Igitur, Divagations, Un coup de des*. Ed. Yves Bonnefoy. Paris: Editions Gallimard, 1976.
- Mayer, Peter. "Some Remarks Concerning the Classification of the Visual in Literature". *Dada/Surrealism* 12, 1983. 5-13.
- Mitchell, W.J.T. *Iconology: Image, Text, Ideology*. Chicago: University of Chicago Press, 1987.
- Mon, Franz. *Gesammelte Texte I: Essays*. Berlin: Gerhard Wolf Janus Press, 1994.
- Peterson, Mary A., Lynn Nadel, Paul Bloom, and Merrill F. Garrett. "Space and Language". *Space and Language*. Ed. Mary A. Peterson, et al. Cambridge: MIT Press, 1999, pp.554-555.
- Pilkington, Adrian. *Poetic Effects: a Relevance Theory Perspective*. Amsterdam and Philadelphia: John Benjamins Pub Co, 2000.
- "Portée." *Larousse English-French, French-English Dictionary*. 1971.
- Rayner, Keith. "Eye Movements in Reading and Information Processing: 20 Years of Research". *Psychological Bulletin*, Vol. 124, No. 3 (1998): 372-422.
- Tsur, Reuven. "Aspects of Cognitive Poetics" on "Cogprints" website, <http://cogprints.esc.soton.ac.uk/archive/00003239/>.
- Underwood, Geoffrey and John Everatt. "The Role of Eye Movements in Reading: Some Limitations of the Eye-Mind Assumption". *The Role of Eye Movements in Perceptual Processes*. Ed. E. Chekaluk and K.R. Llewellyn . n.p. Elsevier Science Publishers B.V, 1992. 111-167.
- Saint-Martin, Fernande. *Semiotics of Visual Poetry*. Bloomington: Indiana University Press, 1990.
- Vos, Eric. "On Concrete Poetry and a 'Classification of the Visual in Literature'". *Verbal/Visual Crossings 1880-1980*. Ed. Theo D'haen. Amsterdam: Rodopi, 1990. 241-280.