

Disjunctive Pronouns: On Multimodal Analysis of Digital Poetry*

METTE-MARIE ZACHER SØRENSEN

Digital poems are genre-mixing entities that raise questions about how to discuss boundaries and combinations of genres, media, and art forms. When is a boundary stable, and how can mixtures be described and analyzed? It is not only relevant to *describe* heterogeneous media configurations and define what they are, but also to ask what they contribute to when it comes to meaning, content and semantics. In this regard, I will use the theory and conceptual framework of Swedish media theorist Lars Elleström (2010) regarding multimodality and mixed modalities. I will argue for the productive use of this conceptual framework in analysis, thus avoiding way too complex terminology, image/text dichotomies, etc. Furthermore, I will define when a work is what I call distinctly multimodal and should therefore be analyzed using these multimodal analytical tools. In the following, I will provide a brief introduction to intermediality theory and, consequently, argue for the fruitfulness of multimodal analysis of digital poetry.

“Intermediality” is a concept with a complex history that, among other things, builds upon the so-called interart studies, where different art forms are compared and examined to understand how they fundamentally differ from each other. One of the most influential historical theorists within interart studies is Gotthold Ephraim Lessing, who reflects on the differences between the two art forms in his work “Laocoön or On the Borders of Painting and Poetry” from 1766. He concludes that poetry (literature) and art (primarily painting) are fundamentally different. Poetry is temporal, whereas art is spatial. Therefore, poetry can describe actions, while art can only depict objects. Art can only depict the same space at once, while poetry can have a plot by allowing different parts to develop alongside each other in a progression.

The concept of intermediality today signals, through its prefix “inter,” that there is something between two media – and that these two therefore differ from each other. In the article “Border Talks” from 2010, Irina Rajewsky describes how one can avoid viewing two entities (media) as essences even though there is an “inter” between them. Media (here understood as “art forms”) and their definitions and boundaries are historically contingent and discursively conditioned. However, they become so thoroughly stabilized that we can still talk about our intuitive understanding of what the different media encompass and how they differ from each other. So even if we eliminate the essentialist aspect, we can still discuss similarities and differences between different media (art forms).

Rajewsky has established a conceptual distinction between media combinations and intermedial references. Media combinations are the result of a process where at least two conventional media forms are combined (Rajewsky, 2005 51–52). Media combinations can be integrated to such an extent that they become independent art forms (for example, films that combine visuals, sound, and storytelling, or operas that blend music, text, and scenography) where their multimodal basis is their specificity. On the other hand, if there is an intermedial reference, it does not involve a combination of multiple media, but rather a process where a singular medium refers to other media by acting “as if” it were a different medium – hence the imitation. A conventional and distinct media type can

thematize, produce, or specifically imitate other media types, but through the use of its own media-specific qualities. For example, one can think of literature that uses cinematic techniques such as cuts, zooms, fading, etc. Regarding media combinations, Rajewsky herself asks:

[O]ne might ask to what extent, in the case of so-called intermedia – including for example visual poetry and corporate logos – one can in fact speak of a combination of different medial forms of articulation, since the constitutive medial forms become *quasi* inseparable. This extreme outer pole of media combinations concerns phenomena in which individual media or their material manifestations – such as word and image – become inextricably bound to, or even “merged” with, one another, and as such are simultaneously and oscillatingly present (Rajewsky 2005, 52)

Rajewsky argues that visual poetry is an extreme form of media combination where words and images are inseparable. However, if we take an emblematic example of visual poetry/concrete poetry like Reinhard Döhl’s “Apfel” from 1965, it becomes clear that one should perhaps apply Rajewsky’s distinctions in a different way by considering visual poetry as an imitation rather than a combination of media.



The work consists of the word “Apfel”, repeated until a formation of words depicting an apple is formed. Among all the words, there is a single “Wurm”, a small worm that has gnawed its way into the apple. The reason I deviate from Rajewsky’s differentiation is that the work does nothing but use the art form of poetry: letters on a page. The difference is that it is not the medium’s range of materials that has changed, but the mode of reception that has changed so that there is a registration of the visual expression of the words rather than a semantic decoding. Words are always visual, but here we have words that act “as if” they were a picture – hence the imitation. Apfel clearly does not meet Lessing’s old definition of poetry. The work is not time-based, but rather spatial – you look at it as you look at a picture. If I, along with Rajewsky, insist that the definition and delimitation of art forms from other art forms are always contingent and discursively conditioned, then we see here an interesting example of a work that forces the boundaries between established art forms to become elastic, or one could also argue that a work like Apfel cements a boundary between poetry and visual art by insisting on saying: “I am a poem and I use the materials that poems have at their disposal – even though I cling right up to the “boundary” of the domain of visual art”. In the end, it becomes a question of institutionally imposed labels: the designation of a material as a poem or a picture.

Apple is a minimalist work – also at the level of intermediality. Because in the following, we will look at an example of digital poetry that has a completely different form of complexity and that combines freely with its use of words, sound, formations, and movement. Lars Elleström has systematized a characterization of media through their modal composition. In the following, I will explain his theory and, at the same time, reflect on the theory as an analytical tool, through its use in a concrete example of analysis.

Intermedial relations and modalities of media

Lars Elleström argues that it is a problem for intermediality research that there has not been a real effort to investigate what a medium is. He points out that media intersect more than they distinguish themselves from each other. One may never stop talking about media boundaries, but his theory is a genuine attempt to transcend them. Elleström aligns himself with W.J.T. Mitchell and his dictum that “all media are mixed media” (2005), but firstly Elleström does not see the historical relations between media as power struggles in the same way as Mitchell, and secondly, he incorporates modality theory. Elleström starts from the bottom, so to speak. Instead of looking at existing media and what they consist of, he examines the existence of basic modalities and interrogate how they are mixed in individual media. The underlying assumption is that all media share a limited number of modalities: semiotic, spatiotemporal, material, and sensory modalities. This conceptual framework makes it impossible, for example, to think of the relation between the verbal and the visual as a dichotomy as it has traditionally been conceived in interart studies (as exemplified by Lessing’s *Laocoön*). Elleström describes how it is assumed that the materiality of literature oscillates between materiality and immateriality, but he emphasizes that the category of materiality itself is unsustainable because it does not consider the fact that different art forms consist of completely different materials, such as language, sounds, and even the immaterial. Sounds are closely related to the sensory, and although language is also a kind of sound, it must primarily be understood in a semiotic context. Elleström further points out that the most common misunderstanding in intermedial comparisons is to confuse the visual with the iconic. The visual pertains to using a particular sense, while the iconic is semiosis based on similarity (which may not necessarily be visible).

Furthermore, Elleström establishes a distinction between basic media and qualified media. For example, a pop song is a qualified medium that includes two basic media: verbal text and organized non-verbal sound. One can discuss whether it is possible to define digital poetry as a qualified medium. Due to the diversity of works, I will refer to them as mixtures of basic media and not attempt to solidify the boundaries for a definition of digital poetry as a qualified medium. I will use Elleström’s model as a starting point for analysis. This means that I not only consider the line of thinking as a reasonable approach to genre and conceptual clarification, but also attempt to show how the model can be used as an analytical tool. I will use the model in a work analysis where I continuously define the different categories and examine how they can open up an understanding of the work’s strategies that go beyond a definition of modes and their mixtures, by also pointing out the effect of these strategies.

Analysis of Digital Poem

The Swedish multidisciplinary artist Johannes Heldén released the digital text work “*Väljarna/Elect*” in 2008. It is a work that combines images, sound, text, movement, and interaction. When the work is activated, it starts with white words on a black background. It is a small amount of text without linear fixation, consisting of the words: “They do not pursue us in dreams // flickering light // distant lights // for the first time real horror // the flocks of birds // move towards darkness // the distant lights¹.”

The formation of words on the page functions as a kind of enjambement with meaningful effects. There are long spaces between the words and insertions on the page: the English phrase “distant lights” breaks with the Swedish text, creating an abrupt reading experience and a sense of fragmented uncertainty. The text is time-bound, so there is not enough time to dwell on it before it fades away. Then, a rectangular frame with pale gray, dappled clouds slowly spread and reveals a black, worn-out tree. It is digitally rendered but naturalistic in its expression. It sways ever so slowly. A drone-like electronic music quietly emerges, creating a somewhat unsettling, threatening atmosphere along with the tree. Above the tree, in a kind of mirroring formation, the shadow of writing

appears, but it is unreadable. After a few seconds, silhouettes of birds begin to fly out between the branches of the tree.



Väljarna/ Elect

One quickly discovers that it is possible to “catch” the birds with the cursor, and then they disappear in favor of a small piece of text that appears below the tree with white letters on a black background. Gradually, one can also click on the text above the tree to make it readable. The order in which the text appears is random because the relationship between each piece and a bird is arbitrary. The amount of text is finite, that is, limited, but there are individual snippets of text that can only stand one at a time, so your access to the work will always be particular in the sense that some text disappears when you want something else to appear. It is a parallel rather than a hierarchical linear expression, and it creates an ambiguity that is reinforced by an ambiguity in the narrative voice. There are truncated sentences with an “I,” a “we,” a “he,” and as mentioned, English statements mixed with Swedish here and there. There are different levels of epistemic competence, for example, an extremely self-referential narrator in a single sequence called “I know everything that can ever happen.” The random order of acquisition entails an uncertainty in the utterance of pronouns. It is not known to whom “they” refer when taken out of context, “They do not pursue us in dreams,” “The distant lights,” and “It”: “It grows quickly now, twines around his legs, knocks him down.” The: “I see it in the corner of my eye and it scares me senseless.” A snippet of a sentence can be: “Others grow too fast and die of exhaustion,” but who are these “others” that the comparison assumes? One lacks the preceding elements that the existence of a comparison necessarily designates. Pronouns (or shifters in a broader sense) can be incomprehensible when standing alone. However, the polyphonic and hybrid text can be said to be categorizable into thematic clusters and different referential levels (i.e., from the self-referential to the referring narrative). The work revolves thematically around the forest and horror: “And he rages, sinks into the ground,” “And stumbles out into the tall grass,” “the voice on the saw blade,” “body parts in the mountainside,” “the knife in the inner pocket,” “drowned in mud,” “all the ice crystals” “the branches of the trees,” “the darkening forest.”

There is an apocalyptic atmosphere that is explicitly expressed when the word “fimbulwinter” appears: Fimbulwinter is in Norse mythology the three-year-long ice winter that heralds Ragnarok. This theme plays exemplarily together with the work’s tree, clouds, birds, and music. But the eeriness in the text moves into other environments, which at the same time are mood breaks because of sharp scene changes when “it” enters an apartment. Thus, environmental breaks are established by moving

from room to room (from a cabin in the woods to an apartment). Other spaces are constituted in a whole, but surreal ways when, for example, a “laughing monkey” appears in writing. It is an absolute break with the Nordic forest with spruces, ice, and abandoned buildings. It is an anthropomorphization to make the monkey smile, and it becomes cute, but also demonic with its smile, so one can almost hear evil monkey screams on an imaginary auditory level that is isolated from the sound image on the material level of the work.

Elleström’s four modalities used in a work analysis

Elleström operates with four base modalities: material, sensory, spatio-temporal, and semiotic modalities. The material modality is defined by Elleström as the medium’s latent physical interface. This means whether the medium’s material is 2D or 3D, whether it is human body, whether it is static or not, whether it is sound waves or laser projections. In practice, it is impossible to separate, but still important to distinguish between the material and its perception. “Väljarna” as a work has a materiality with moving images and sound waves. The screen is two-dimensional – it has width and height – but already through interaction, the mouse function, etc., the appropriation extends into a spatial and temporal dimension. This concerns very basic phenomenological experiences of the computer medium.

The sensory modality is the physical and mental action of perceiving a medium’s interface with one’s senses (Elleström, 2010, 17). Normally we talk about the five human senses: seeing, hearing, feeling, tasting, smelling, but according to Elleström, three more must be mentioned: sensory data originating from objects, phenomena, etc. that cannot be encapsulated in isolation without a perceiving and interpreting agent. These three are internal creations: feeling a sculpture even though seeing it, hearing music and speech, it has much to do with an internal balance, feeling the rhythm in the body, etc. One imagines images when reading, even though they differ from the appearance of letters on the page. Think about how in “Väljarna” I see and hear the cute and demonic monkey at the same time independently of (or rather as a contrast to) the work’s materiality (its flatness, sound, and images).

As mentioned, the four modalities are difficult to separate, and examples of the sensory modality will appear in the following when I analyze the music and the ways in which the tree is temporal, their ways of creating atmosphere and space, as well as the text’s way of establishing space, and soundscapes on the material level of the work. In the interaction with the work, the birds become tools – buttons that can be pressed to bring forth sentences instead of images in themselves. At the same time, they are of course also bird pictures that, together with the black tree, create atmosphere and evoke associations and dissociations along with the words. But this is a good example of how the interface and the work, so to speak, have merged – in the digital realm, links and buttons are symbolic signs – and with the birds, we see how what needs to be done with the work is a part of the work itself.

The constitution of time and space at different levels

The spatiotemporal modality is an important tool for analyzing “Väljarna”. Elleström describes how cognitive factors cause all media to be presented with both spatial and temporal qualities. Spatiotemporal perception can be said to consist of four dimensions: width, height, depth, and time. A photograph’s physical interface only has two dimensions: width and height. A sculpture has three material dimensions, all spatial: width, height, and depth. Dance has all four dimensions – a dance performance has a beginning, an extension, and an end – and therefore exists in time, while a photograph just exists. You don’t miss anything if you close your eyes for a moment – you miss something with a dance performance. Media that lack the fourth dimension, time, are static. Media that incorporate the dimension of time in their physical manifestation can have:

- fixed sequentiality (moving images and recorded music, such as the music in *Väljarna* that cannot be manipulated)
- partially fixed sequentiality (e.g. hypertext and video game music – this is what is happening with the text in *Väljarna* where there is interaction, but the text moves independently of you in chunks).
- unfixed sequentiality (mobile sculptures, improvised music, live, televised performance).

Elleström describes how media can be spatial, and the most fundamental form of spatiality is found at the level of material modality: we have physical width, height, and depth. However, this is not the only form of spatiality because he believes that our perception largely functions spatially. Abstract concepts and experiences of time also have spatial characteristics (Elleström 2010, 20). So the virtual space is also important: images and photographs are flat but give the impression of depth. Narratives in text provide a “real” but virtual world that one can navigate. Just as the concept of virtual space, there must be a discussion about virtual time. Some media have physical interfaces that are not temporal, but all media are realized in time. One can think of a static image with motifs indicating a time, for example, a before and an after the current configuration. Written text on paper obviously does not possess any temporal quality other than what you yourself evoke and control, unless one wants to speak of rhythm as indicating tempo, but of course, text on the referential level can possess temporality – narration necessarily unfolds in time. In “*Väljarna*”, the temporality of the text is complex. On its material level, it has a partially fixed sequentiality – you are not controlled in your reading order of lines, but on the other hand, you do not determine each time what you want to read and what remains. You barely have time to register the sentence “Body Parts in the Rock Wall” before it disappears again. This contributes to uncanniness on a sensory level. The semantics of words are thus influenced by the way we acquire them. Like when you see a short clip of an image in a film and do not have the opportunity to stare until you have understood, dismissed, or mentally disarmed the motif. In some sequences in “*Väljarna*”, the writing appears letter by letter, as if seeing the writing as it is being written. Thus, one is fixed in the sequentiality – your reading speed is controlled. There is a special kind of temporality on the material level. In one case, there is a sort of coincidence between this level and the referential one. The sentence: “it’s growing fast now, winding around his legs, tearing him down” moves at a writing speed, approximately in my reading order. It feels as if the movement on both levels is in sync because they both have progress. “*Väljarna*” deals with the phenomenon of time on a referential, thematic level with some abstract and aphoristic sequences such as: “everything exists infinitely” and “the curve of time is soft”. A statement about a soft curve of time can seem like a colossal shift in levels if one has just read about rust on a saw blade. Any associative, metonymic or thematic connections between statements are nullified when the display of sentences has a random sequentiality. The statement about a soft curve of time induces dizziness, a distortion of the metaphor of time as a curve. Can one then pull on the curve, stretch it like a rubber band, or is it soft in a way that allows one to make holes in it? The image must stand alone because you do not dare to trust a connection to other semantic statements in the work. The appropriation of words, with the abrupt sequences as mentioned, can give the sense of lacking intentionality, but conversely, the work plays with other mood-enhancing techniques that establish a sophisticated form of appeal. I am thinking here of the other basic media that constitute the work: music and images. Visually, the work is dominated by the digital tree that moves ever so slowly. One could say that it is a stagnant, almost cyclical form of movement fixed in its sequentiality (if one can even speak of such). In that sense, the tree exists on the same temporal modal level as the auditory mood creator of the work: the music. It is, of course, temporal, but it resists its own temporality and establishes an atmospheric underlying mood with its minimal, drone-like sound. Rather than a temporal progression, it is a kind of trembling, and a potential sense of progression is constantly annulled because in the background of the sound image, a careless oscillation between two deep tones is hinted at. In a sequence in the text, there may be a reference to the work’s own music: “like believing that the noise does not hold any melodies, because it does. They are there, and they are

more than can be counted.” The music precisely allows for a multitude of virtual spatial openings and progressions. The minimal is simultaneously an opening and a closure. The interaction between the tree and the sound creates a specific mood. Even though the sound does not mimic the “sound of a tree,” they would not be able to do without each other. The reason why there is always music for silent films is that moving images without sound appear extremely ghostly. Movement is friction and friction produces sound. The fact that the artwork actually has a slightly ghostly atmosphere is due to the deep underlying tone of the music, the gloomy, gray-toned colors, and parts of the text. When I analyze how the music and images function as emotional sutures – as something that establishes a kind of coherence and address despite the rejecting gesture in the text, the question arises of whether these can be said to “fill in the gaps” in the text. In “Väljarna”, there is the following sentence: “Now, now it enters my apartment and moves towards me.” When we suddenly find ourselves in an apartment we have never been in before and the sentence is torn out of any context, how do we contextualize the “it” in the sentence? When the text consists of disjunctive pronouns, can image and sound create a context that perhaps does not close, but delimits the space of meaning? A space of meaning that otherwise stands wide open with the use of, for example, “it” or “that”.

The space of the sky is a completely different place – The semiotic modality

In *Väljarna*, I also find examples of words “escaping” from image and music. To nuance these effects, I will briefly introduce the last of Elleström’s modalities: the semiotic modality. This modality is divided into Peirce’s well-known triad, which distinguishes between symbolic, iconic, and indexical signs. Elleström emphasizes that the semiotic nature of media is immensely complex, but there are inevitably basic semiotic differences between, for example, written text and a moving image (Elleström 2020, 22-23). When the words in “*Väljarna*” create spatial environments independent of the image and music, we see a difference between the effects of the iconic and the symbolic in reception. It is interesting how the words of the work (as symbolic signs) interact with the visual and the auditory, but can also create their own distinct visual spaces as a contrast: “See a pale yellow sun rise, a supernova explode, and the pupil expand”; from the realistic, picturesque scene with a pale yellow sun rising, to a supernova (a star much, much larger than the sun) exploding, and then down to the relatively small pupil expanding slightly. It is a small reaction to a supernova explosion. One is the colossal leaps in sizes for expansion here in the imaginary space, but imagine if, after the text about the supernova, you jump to a self-referential text sequence about the birds that we are trying to catch. From abstract, gigantic events in the sphere outside the space of the sky, to your movement right now with a computer cursor capturing a small, black pixel bird that produces a word. The birds become buttons while also being mood-setting birds, because they constantly oscillate between being iconic and symbolic signs. But, of course, they are not symbolic in the same way as words. When Elleström rejects the verbal-visual dichotomy, it is because the verbal has a “body” that is visual or auditory – but it does not mean that everything becomes the same. Text is still text with its own effects – and image and sound as well. When I examine configurations of space in the work, the point is precisely that there is a difference in the auditory, semantic, and visually figurative representation of it. The small, black letters on a grayish background send the recipient (imaginary) visual references to supernova explosions, and we must once again think of the smiling monkey who visually and audibly places itself in an imaginary place. It seems to be an intended effect in the present work that we both scratch the surface of the “concrete” material by interacting with the flying birds and physically sense the rapidly moving letters *and* get thrown into space and apartments, as well as having the imaginary sound effects laugh alongside the actual digital drone sound.

Conclusion

The modal composition is located at the level of meaning formation. With Elleström’s model, I have a conceptual framework and a conceptual mindset that I can use in analyzing genre-challeng-

ing works. I don't expect the genre of digital poetry to stabilize and become a qualified medium in the same way as, for example, film. With digital poems (and much other contemporary art), the bag of modal blocks is shaken every time. While the modal composition of well-established art forms can be said to largely reside at the level of established conventions, distinctly multimodal works have modal compositions that include distinct and "new" elements in the appropriation of the work, which have a significant impact on meaning formation and therefore require a non-trivial effort from the recipient (Aarseth, 1997, 1). The modalities must therefore be analyzed as more and different than just the genealogical characteristics of the work's genre because they have not (yet) been internalized. In other words, the modal composition belongs to the level of meaning formation. The modalities and their effects are functionalized and externalized in the sense that in encountering the work, they are experienced as if they are on the "outside." They are experienced as more than just 'being there'. Therefore, it makes sense to integrate a sensitivity to this composition and its effects as part of one's analysis.

Aarhus University, Denmark

Notes

* In Memory of Lars Elleström (1960-2021).

¹ My translation from Swedish.

Works Cited

- Aarseth, Espen J. *Cybertext- Perspectives on Ergodic Literature*, The John Hopkins University Press, 1997.
- Lessing, Gotthold Ephraim, *Laokoön oder Über die Grenzen der Malerei und Poesie*, Projekt Gutenberg-De, (1766).
- Rajewsky, Irina O. "Intermediality, Intertextuality, and Remediation: A Literary Perspective on Intermediality" in *Intermedialité / Intermediality*, (6), 43-64, 2005.
- Rajewsky, Irina O., "Border Talks: The Problematic Status of Media Borders in the Current Debate about Intermediality" in Elleström, Lars (red.), *Media Borders, Multimodality and Intermediality*, Palgrave Macmillan, 2010.
- Mitchell, W. J. T., "There are No Visual Media" in *Journal of Visual Culture*, Vol. 4, No. 2, 2005.