

TNA ACTIVITY REPORT

PROJECT TITLE: Criss-crossing Novel Illustrations and DALL-E-
generated Images: toward a data-rich ekphrasis

Author: Roxana Patras

Current position: Senior Researcher I

Affiliation: “Alexandru Ioan Cuza” University of Iași

Host institution: University of Trier

Mentor(s): Prof. Dr. Christof Schöch

Period of stay: 14.06.2023 - 19.07.2023

Brief introduction with the presentation of the research questions:

My grant proposal has been built upon a theoretical frame that defines *ekphrasis* as an inter-semiotic “conductor”. Notwithstanding this oversimplified take on visual poetics, the approach would still enable us to study how visual cues transgress from texts to images and viceversa.

The main research question was if the visual cues in textual bits such as portraits and descriptions are strong enough to produce, via text-to-image tools such as Dalle (<https://openai.com/dall-e-2/>), <https://dallery.gallery/dall-e-ai-guide-faq/>), Stable Diffusion (<https://stablediffusionweb.com/#demo>) or others, images. If the output considered relevant in terms of visuality, then the adjacent research question was if the AI-generated images could be compared (and on what grounds) with the illustrations of the original texts.

Accordingly, the first objective was to search deeper into the analogical relationship text-image and to carry out a methodological experiment on both types of transformations: a.



image-to-text, as in the description of pictures/ artifacts, which has always provided the most commonplace example of *ekphrasis*; b. text-to-image, as in the process of illustration, developing from the old Egyptian *Book of Two Ways* to our times' comic books.

The second objective was to document: a. the theory of *ekphrasis* from Murray Krieger's seminal book *Ekphrasis: the illusion of the natural sign* (1923) to newer contributions from stylistics and iconology (Maria Eirini Panagiotidou, Ruth Webb, Peter Wagner, W.J.T. Mitchell, and others); b. the theory of illustration, with a stress on 19th-century illustrated novels and press.

The third objective was to evaluate resources for the computational study of literature in lesser-used-languages (with a view to the broader integration of LUL literary corpora in multilingual investigations) and to carry out a methodological experiment on what might be called a "backward restoration" of books' materiality. In my opinion, after the heyday of massive digitization projects, a reconsideration of the original artifact is timely and necessary in CLS, chiefly when we evaluate the gains and losses of digitizing bigger or smaller literary collections such as ELTeC, DRACOR, Hai-Ro, Pop-Lite (<https://www.dracor.org/>, <https://distantreading.github.io/ELTeC/>, <https://poplite.org/>, <https://proiectulbrancusihairo.wordpress.com/>). As recent research straightforwardly expresses or only suggests, the textualism's authoritarian legacy has driven to the prevalence of text analysis over image analysis (Manovich 2012) as well as to a nostalgia for what the text could signal in terms of "embodiment" or, better said, embodied reception: indeed, an attempt at battering down the towers of textualism can also be grasped in the scholars' recent turn toward synesthesia instances, such as expressing values of sounds, touching or smell through various textual devices (see Piper 2023, Guhr and Algee-Hewitt 2023)

Methodological plan:

My first idea was based on a very simple triple step: 1. to annotate texts or to spot the sequences having some sort of portraiture "density"; 2. to extract text chunks that would retrieve bits/ entire portraits or descriptions; 3. to generate images from extracted fragments. Also, greatly influenced by Genette's theory (Genette 1982) but also by my recent analyses of paratext (such as titles or prefaces), I have hastily considered that whatever the dime illustrations included, their status is "para", thus text-dependent. But the initial plan was not effective for several other reasons that are explained below:

a. It did not properly distinguish, based on vocabulary variety or on narrative functions, between human descriptions (portraits) and other types of descriptions (setting, objects).

b. Even if we accepted that the generally-acknowledged stereotype of dime/ popular novels should ensure "fixity and stability" of word-image relationship, the one-on-one, completely unmediated analogy - based on, say, visually-enhanced vocabularies, including portraiture elements - would be hard to proof further.

c. The annotation raised various kinds of problems: 1. the affective and expression determiners of corporeal elements are not always represent-able (e.g. "sad eyes"); 2. even if we discriminated between the features corresponded in illustrations (e.g. `seg corresp="#Bujor" sameAs="Imagine1">`) and features that are present only in the text, the aggregation of characterization sequences for each name (e.g. `<person xml:id="Bujor"><name>Bujor</name></person>`) in the `<listPerson>` would lead to a robot

portrait. A short note should be made on the fact that aggregation is only possible when entire sentences are annotated, while other relevant descriptive bits - if considered for extraction - should undergo a kind of “reconstruction” and “reordering” in a new sentence matrix; 3. the distinction between @ana="description" and @ana="designation" needs more textual examples in order to be proved as effective. In any case, problematizing the correspondence or the autonomy of illustration from/ with the surrounding text (see below) means to decide between the status of text wrapped around illustrations and the status of illustrations embedded within the text, and only then to prop annotation accordingly.

d. My research focused on the semiotic correspondences between the texts of 19th-century Romanian dime novels, their illustrations, and AI-generated images. Therefore, closer attention should have been paid to the cultural context of production and reception of images, in our case, dime novel illustrations. Beside text-to-image conversions, a culturally-informed research should have addressed the type of “gaze” that dime illustrations would elicit, more precisely “perplexity” and “suppression of the multiplicity of the text” as readers’ possible responses, the kinds of reaction that would already be abhorred by high-literature authors (Thomas 2017, 54). Based on data regarding the Romanians’ literacy level in the period when these dime novels were published (between 1870 and 1912, the literary rate only doubled from 20% to 40%), we should have probably asked first how would a typical (that is, illiterate or semi-literate) reader *gaze* at those novel illustrations? Was the “multiplicity of text” still being “suppressed” or, on the contrary, enhanced? Because I also discovered that some of the library copies, presumably preserved in better condition, would have the illustrations clipped out from the body of the text, the subsequent question - though not directly related to the aims of the present exploration as many others formulated throughout this report - is if the circulation of clipped illustrations on another route than the core text could have influenced the illustrators’ choices when picturing novel scenes.

e. Even though the aim of my exploration was initially to produce, via a referential and mimetic relation, “images” from “texts”, the implied (but not overtly declared) meaning of image was “illustration”. Thus, what I wanted to produce, via AI text-to-image platforms, were novel illustrations. While illustrations’ histories and archives have recently flourished in established literary traditions (see for instance English archives of 19th-cent illustrations) - that is, traditions with established printing protocols and established literary markets -, there is still a low interest for the illustrations of dime novels, because they have always been considered base and cheap, the oeuvre of anonymous authors and the spawn of recycled blueprints. For peripheral literatures such as the Romanian, the phenomenon is barely considered as a proxy of caricature or satyric press (Oprescu 1945, 95; Bădicioiu 2022, 171-183). Having mentioned all these matters, a more informed perspective on the theory of illustrations is required.

Steering clear from the aforementioned ineffectiveness, I tried to eschew textualist methods and re-approach the research question(s) by blending new narratology (a multimodal definition of the novel), iconology (W.T.J. Mitchell’s vitalist approach to image reception, boiling down the fact that images/ pictures want from us something that they lack), recent contributions on visual digital archives (Julia Thomas’s view that illustrations might own features that are “invisible” in print form but that can tun visible by digitisation processes), and a bit of neuropsychology (Ardilla *et alii*’s theory of the illiterate brain and of its particular way of decoding images). Accordingly, the exploration of text-to-image conversion

turned into a reader-response laboratory. A deeper consideration of the Romanian dime novels' context of reception (literacy level, the semi-literate's or the illiterate's gaze) has led me into a necessary differentiation among 3 types of reading elicited by dime novels containing illustrations: a. reading images; b. reading texts, c. reading images+texts. As I explained above, the success of the 3 types of readings depends on the literacy of the reader, thus on how fast each semiotic system is decoded. To the point, for a semi-literate or illiterate reader such as the 19th-cent Romanian reader, illustrations are definitely read/ decoded faster than the text. This means that illustrations acquire a sort of autonomy (they turn into "totems", with lives of their own, as Mitchell shows) and do not function as textual "suppressors" (as Thomas claims) but as incentives for a future reading of the text. Apart from the reader's speed in decoding the two main systems of signs (letters and images), the success of the 3 types of readings also depends on getting the right visual "anchors" in the images, as on getting the right textual "anchors" in the body of the text. But, if the text avails of a typeset and of a partitioned structure as anchors, what are the zones of anchorage in illustrations, especially in the illustration that aims at catching the sight of a semi-literate or illiterate readership?

Description of the research visit and its outcomes:

The tasks and activities carried out during the research visit organized according to 2 scenarios:

The first scenario involved the evaluation of tools and resources for text-to-image conversion, the translation from French to Romanian and the augmentation of a body-related vocabulary, the annotation of character-related sentences in 10 dime novels, the compilation of an archive of novel illustrations, and the participation in a CLS conference. Because at the moment of the grant application, DALL-E was still allowing free credits, I assumed that the platform would also provide support during the research visit. Since it was not the case, I resolved to work with Stable Diffusion (SD) and with ChatGPT3 instead. As illustrated in the excerpts below, even with a Lesser-Used-Language(s) input (Romanian, sometimes archaically spelled, in my case), ChatGPT3 performed quite well for tasks such as summarization, excerpting, reported speech, keyboarding and so forth.

The experiments developed as follows:

1. EXTRACTION> I extracted: a. novel scenes that contained enough descriptive data (ex. Introductions/ exordium); and b. novel scenes that are correspondent to illustrations. I also tried to extract portraits from entire chapters with ChatGPT by prompting it "select a description from the fragment" "extract textual sequences that refer to persons/ that contain at least two features of persons/ that characterize persons".

2. SELECTION> I selected the lengthiest portraits and prompted ChatGPT "*illustrate* the following textual sequence...". Very interesting, the answer appeals to the user's imagination and ability to visualize a narrated scene: "I apologize, but as a text-based AI model, I'm unable to directly provide visual illustrations. However, you can imagine the main character based on the descriptions provided. In your imagination, visualize..."

3. ENHANCING VISUALITY with AI> I analyzed the output of ChatGPT3: generally, it can be regarded as a visually enhanced version of the original sequences, produced through summarization and narrativization. But I also tried to identify the visual

cues introduced by the AI so as to “appeal” the user’s “imagination” and capacity to “vizualize”.

4. ENHANCING VISUALITY with additional textual data> I returned to the original text, chose randomly more portrait features from other chapters of the novel (other than the one fed in ChatGPT3), then added these features to the ChatGPT3 output. Apparently, postural information and other body segments and body-related objects are necessary.

5. IMAGE GENERATION: I prompted SD with “generate a (landscape, portrait of a woman, drawing room *picture*) in (19th-century illustration, Max Ernst, Damier, Hogarth, etc.) *style of (sequence of text to be inserted...)*” and then I analyzed the granularity of the output. In the majority of cases, the main problem appears in the representation of faces, which are drawn as a blurred oval because “face information” is lacking from the input or is not sufficiently “iconic” (expression hints). In fact, scholars have already noticed the actual poverty of portraiture in written texts, thus the unreliability of textual information related to portraits (see above my comments on the vocabulary for description detection): „Peter Mendelsund has remarked on how little information readers are given in texts about what fictional characters look like. ‘It is precisely what the text does not elucidate’, he comments, ‘that becomes an invitation to our imaginations.’ To emphasise this point, Mendelsund’s article in *The Paris Review* includes numerous illustrations of figures with their faces and heads erased or chopped off. The problem with illustrations, as Mendelsund’s pictorial beheadings suggest, is that, by showing details that the texts do not, they negate this imaginative interaction and instead direct and manipulate our interpretations, ‘closing off’ the plurality of the text. This is precisely the reason why illustration was so frequently criticised in the nineteenth century“ (Julia Thomas 2017, 54). The face problem could be probably disentangled if we turned toward the original illustrations of the dime novels, and tried to segment the images: Mandelsund’s idea to erase or chop off heads and faces from illustrations could turn out interesting results in terms of image hierarchy.

Below, I paste in a sample of ChatGPT3 conversation:

“user

extract the descriptions from: "<p>Noaptea furtunoasă,</p>

```
<p>Era pe timpul fanarioșilor. Toată țara gemea. Pământul  
României pe fiecare zi, pe fiecare ceas se  
păta de sânge nevinovat, și aerul său  
răsuna de gemetele victimelor, sau se înegrea  
de mârșave crime.</p>
```

```
<p>Femeia frumoasă, copila nevinovată trebuiau să fie  
jertfe poftelor nesățioase a veneticilor; banul  
muncitorului, averea românului trebuiau  
să intre fără murmură în marea pungă a  
Domnitorului și ca răsplată pentru toate  
acestea, locuitorii acestei țări trebuiau să  
îndure mizeria, moartea sau chinul fără  
marginii!</p>
```

```
<p>Era timp de iarnă, două zile ninsese de a rândul și  
acum o furtună mare împrăștia prin aer fulgii
```

de zăpadă.</p>
 <p>Întunericul nopții era bătut de albeța ce răspândea neaua. Stelele păreau ca niște cuie de aur înfipte în cer. Pe singuraticile străzi ale Bucureștilor un om se vedea traversând podul Mogoșoaiei. El era urmat de alți patru.</p>
 <p>După fața lui aproape înghețată de frig, după cizmele de lac și după blana cea mare de vulpe cu care era învelit, trebuia să fie un om cu punga groasă, prin urmare un străin, adică un fanariot.</p>
 <p>– Ei! Vutira, pe unde trebuie să mergem? întrebă el pe unul din cei patru care îl urmau.</p>
 <p>– Drept înainte, stăpâne, răspuse cel mai înalt dintre ei.</p>
 <p>– Și mai avem mult drum?</p>
 <p>– Ca vreo zece minute.</p>
 <p>– Îmi pare rău că am plecat pe jos.</p>
 <p>– V-am spus aceasta, stăpâne.</p>
 <p>– Cu toate astea am plăcere să fiu original în noaptea acesta. Mi s-a urât a sta trântit tot în trăsură, înghețând de frig, intrând apoi într-o mică căsuță, stând pe un pat de paie și încălzindu-ți mațele la căldura unei inimi, cred că e ceva plăcut măi Vutira ?</p>
 <p>– E mai mult decât plăcut, stăpâne.</p>
 <p>– Dupe cum văd eu, tu nu pari a fi un om prost.</p>
 <p>– După cum va crede Prea-Înalta-Strălucita-Lumina-Voastră!</p>
 <p>Se făcu tăcere. Câte și cinci urmară drumul înainte, înfruntând zăpada.</p>
 <p>– În căsuța asta, zise apoi Vutira.</p>
 <p>Câte și cinci intrară înăuntru.</p>
 <p>Odaia era puțin încălzită de focul ce licărea într-o sobă de fier. Mobilierul era de lux, dar învechit. El se compunea din două fotolii, patru scaune, o oglindă mare, două dulapuri și două paturi. Pe masa din fața oglinzii se vedeau câteva cărți și vreo două mici statui de artă.</p>

<p>Totul arăta că cei ce locuiau această odăiță, se bucurau odată de-o avere frumoasă și că aveau o educațiune mai aleasă.</p>

<p>Două ființe se vedeau însă acum în mica odaie. Una era albă ca iarna, cealaltă tânără și frumoasă ca primăvara. Amândouă aveau vestminte negre și la amândouă ochii erau udați de lacrimi. Nenorocitele femei aveau durerea a pierde, una pe soțul ei, cealaltă pe iubitul ei tată.</p>

<p>Trei zile au trecut numai din seara, când frumoasa Mărie, conduse la ultima locuință pe bătrânu-i tată.</p>

<p>Durerea și disperarea ei nu se pot descrie de până nimănui.</p>

<p>Trei zile trecuse numai și ea slăbise mult, ochii i se stinseseră; totuși, era încă frumoasă.

Trăsurile feței sale erau nobile, aerul ei majestos de patriciană și ochii săi, cu tote că pierduseră vioiciunea, păstrau însă dulceața și acel farmec cântat așa de mult de poeții romantici.</p>

<p>Când cei cinci străini intrară înăuntru, femeile se speriară crezând că sunt hoți. Și oare, nu aveau dreptate?</p>

<p>Tainicul străin, acela, pe care Vutira îl numea: <hi rend="italic">stăpâne</hi>, aruncându-și ochii asupra Măriei, nu putu să nu facă un gest de satisfacțiune. Copila i s-a părut destul de frumoasă.</p>

<p>— Ce voiți, domnilor? întrebă muma, pe când copila se retrase înspăimântată spre colțul patului.</p>

<p>— Ce vrem, răspunse străinul, vei afla mai târziu, mulțumește-te a ști numai că de azi înainte veți fi bogate. Cunosce nenorocirea de care sunteți izbite...</p>

ChatGPT

Here are the descriptions extracted from the given text:

1. "Femeia frumoasă, copila nevinovată trebuiau să fie jertfe poftelor nesățioase a veneticilor" - **The text**

describes a beautiful woman and an innocent girl who were targeted by the insatiable desires of the Venetians.

2. "El era urmat de alți patru." - The main character is a man who is being followed by four individuals.

3. "După fața lui aproape înghețată de frig, după cizmele de lac și după blana cea mare de vulpe cu care era învelit" - The main character is described as having a face almost frozen from the cold, wearing lacquered boots, and wrapped in a large fox fur.

4. "trebuia să fie un om cu punga groasă, prin urmare un străin, adică un fanariot." - Based on his appearance, including his thick purse, the main character is identified as a wealthy stranger, specifically a Fanariot.

5. "Câte și cinci urmară drumul înainte, înfruntând zăpada." - The main character is accompanied by five individuals as they continue their journey, braving the snow.

6. "Odaia era puțin încălzită de focul ce licărea într-o sobă de fier." - The room they enter is slightly heated by a fire burning in an iron stove.

7. "Două ființe se vedeau însă acum în mica odaie. Una era albă ca iarna, cealaltă tânără și frumoasă ca primăvara." - Inside the room, there are two individuals present. One appears pale like winter, while the other is described as young and beautiful like spring.

8. "Amândouă aveau vestminte negre și la amândouă ochii erau udați de lacrimi." - Both individuals in the room are wearing black clothing, and tears fill their eyes.

9. "Trăsurile feței sale erau nobile, aerul ei majestos de patriciană și ochii săi, cu tote că pierduseră vioiciunea, păstrau însă dulceața și acel farmec cântat așa de mult de poeții romantici." - The description refers to the noble features of the young woman's face, her majestic air of a patrician, and her eyes that,

despite losing their liveliness, still retain their sweetness and the charm praised by romantic poets.

10. "Tainicul străin, acela, pe care Vutira îl numea: stăpâne, aruncându-și ochii asupra Măriei, nu putu să nu facă un gest de satisfacțiune. Copila i s-a părut destul de frumoasă." - **The mysterious stranger, referred to as "stăpâne" (master), gazes at Maria and cannot help but express a gesture of satisfaction upon finding her quite beautiful.**

Please note that the descriptions provided are based on the extracted text and may not encompass the entire characterization of the main character.

user

Please extract the description of the main character

ChatGPT

Based on the provided text, **the main character is described as follows:**

- **The main character is a man who is traversing the streets of Bucharest during a winter storm.**
- **He is bundled up in a large fox fur coat to protect himself from the cold.**
- **His face is almost frozen from the frigid weather.**
- **He is wearing lacquered boots.**
- **The main character is identified as a wealthy stranger, specifically a Fanariot, based on his appearance and the assumption that he carries a thick purse.**
- **The man is accompanied by four individuals who are following him.**
- **He is addressed as "stăpâne" (master) by one of his followers, indicating a position of authority or respect.**
- **The main character expresses a desire to be different and original on this particular night, preferring to travel on foot rather than in a carriage.**
- **His presence elicits a sense of satisfaction upon seeing the beautiful young woman named Maria.**

These descriptions provide insights into the physical appearance, social status, and demeanor of the main character in the given text.

user

please provide me with an illustration of this fragment

ChatGPT

I apologize, but as a text-based AI model, I'm unable to directly provide visual illustrations. However, you can imagine the main character based on the descriptions provided.

In your imagination, visualize a man walking through the snowy streets of Bucharest during a winter storm. He is wrapped in a large, luxurious fox fur coat, which adds an air of elegance and wealth to his appearance. His face is nearly frozen, with a hint of determination or seriousness in his expression. He wears polished lacquered boots, emphasizing his attention to detail and refined taste.

As he walks, the man is followed by four individuals, suggesting a sense of authority or leadership. You can imagine them trailing behind him, perhaps mirroring his stride or posture. The main character exudes an aura of confidence and control, commanding the attention of those around him.

To add depth to the scene, picture the dark night illuminated by the whiteness of the falling snow. The glow from the streetlights reflects off the snowflakes, creating a soft, ethereal ambiance. This setting enhances the contrast between the main character's dark attire and the surrounding wintry landscape.

Please note that this is a subjective interpretation, and your own visualization may vary based on your personal imagination and artistic style."

With the 2 portraits rephrased by ChatGPT3 (marked in the excerpt above in blue), the free SD generated the following images, some of them with "frozen faces", strangely combining a man's frame with "a fur" and, in the case of the second portrait, adding a fur hat even if the input text does not mention it. "The air of elegance and wealth", "attention to detail and refined taste" is conveyed by a "catwalk" posture (one foot foregrounded, hand in pocket).



Explicație





The second scenario involved the exploratory manual annotation of 50 illustrations of dime novels, the conceptualization and operationalization of categories with a view to an ontology tailored for illustration annotation, documentation and testing of image annotation tools, documentation of authority references (on iconology and ekphrasis theories), experiments on the visual enhancement of textual sequences. This scenario might be considered “retrogressive”, because it travels back, from illustration-to-text, but it might also be considered as visually “enlightening” as it was meant to improve the illustration’s “searchability” and “visibility” (Thomas 2017, 33-65). My main point when trying it out was that not all textual descriptions are visually impacting enough so as to be checked as fit for illustrating. Thus the approach involved identifying the zones of anchorage in dime illustrations, developing a model for the annotation of illustration-anchors, and then evaluating the annotation scheme against the documentation of AI trained models. For instance, the SuperAnnotate tool enables object detection, pose estimation, pixel-precise segmentation, classification and OCR (<https://www.superannotate.com/>).

There are other reasons for testing a retrogressive, slow reading of dime illustrations and novels. Inspired by Ardilla & others’ finding about “the illiterate brain”, the scenario reenacts the illiterate’s or the semi-literate’s perception of the novel’s printed page. In a nutshell, schooled literates and illiterates use not only different neural networks to perform the same task but they differ in terms of visual memory and visual perception, the illiterates being prone to perceive and remember concrete details and concrete objects (Ardilla and others 2010).

The experiments developed as follows:

1. IMAGE CLUSTERING> I prompted SD with clusters of illustrations from the archive I created. The clusters have been organized in various ways: a. thematically/topically (e.g. illustrations of “hajduk heroes” from the HAI-RO corpus of hajduk novels); b. monographically (e.g. illustrations of one novel and, if the case be, of its various editions); c. author-wise (e.g. illustrations of the same author’s novels).

2. RETROGRESSIVE READING> I practiced a sort of “retrogressive reading” by simplifying these images, usually to cartoon form.

3. AI-ruled HIERARCHY> I identified the iconic elements that are kept and in what order of importance (the hierarchy of visual foci).

4. RE-NARRATIVIZATION OF IMAGE> The step concerning the conversion of image to text - the narrativization of image - is not entirely clear in terms of platform options. In any case, textual sequences like the illustrations’ corresponding scenes should be riveted.

5. MANUAL ANNOTATION> I manually annotated the visual “foci” of 50 dime novel illustrations, by assuming that the order of image decoding/ reading might be ruled by 2 kinds of neural patterns: a. the literate or semi-literate will decode images by using the order of typed words (generally, from left to right and top to bottom); b. the illiterate will decode images by paying attention to the their deictic force, practicality and concreteness (generally from bottom to top, attention paid to the foreground elements and to daily use, recognizable objects). The advantage of establishing image hierarchy through manual annotation could be suggested by Julia Thomas’s comments on crowdsourcing

projects such as *First Impressions* that aimed to gather data from users who clicked on the first thing that caught their eye.

Potential Markup Categories:

| Person | | Object | | Attached_text | |
|-------------------|--|-------------------|---|---------------|--|
| Gender | a.male b.female c. invalid | Place | 1. Front 2. Back 3. Right 4. Left | Content | 1. Summary 2. Citation/ Excerpt |
| Dressing | a. pieces of clothing (vest, scarf, coat) b. nude | Size | 1. Big 2. Medium 3. Small | Form | 1. Emphasized words 2. Other markup (punctuation) |
| Posture/ Pose | postural verbs (e.g. "seated") | Shape | 1. Angular 2. Curve; 3. Human (e.g. for shadows/portraits- icons on the wall) | | |
| Body Elements | 1. Hair 2. Face- Cheek 3. Eyes 4. Nose 5. Mouth 6. Breast- Bosom 7. Chest 8. Belly 9. Feet 10.Hands | | | | |
| Attached_Elements | 1. Object (tree, bottle, candle) 2. Person (male, female, invalid) | Attached_Elements | 1. Object (tree, bottle, candle) 2. Person (male, female, invalid) | | |

A note on attached texts or captions: “When applied to illustration, this text derived notion of a caption as a ‘heading’ is literally turned upside down, with captions commonly placed at the foot rather than the head of an image, as our definition of a caption as ‘a title below the image’ implies. The caption is turned upside down in another sense, too. Whereas the objective of a caption in the sense of a textual heading is to explain or encapsulate what the article is about before it is read, the caption of an illustration is often read alongside or even after a viewing of the image. Moreover, the caption of an illustration does not necessarily play an explicatory role. The nature of the interaction between caption and illustration is, in some ways, determined by the genre of text” (Thomas 2017, 44)

Considerations over future work:

The three objectives of the research stay have been achieved. Given the richness of input theoretical insights and the original (if not extravagant) methodological mix, I also tossed around a few other conjectures and must-try(s) that, with the accumulation of more case studies, will be probably drafted as an article:

1. The two transformations (a) *text-to-image (progressive)* and (b) *image-to-text (retrogressive)* imply different strategies and approaches as far as annotation is concerned: (a) *text-to-image* involves the annotation of linear (textual) information and then its clustering according to semantic affinities; (b) *image-to-text* involves the annotation of tabular, stacked (iconic) information and then its clustering according to visual foci.
2. The illustrators of the dime novels also perform a special kind of reading that turns textual, linear information into stacked information (the images). Probably, the illustrators’ reading is indexical and not linear, that is, they take out from the string of words what is important and indicate, through visual techniques, the text’s main *topoi*/ topics. As in the case of the “illiterate brain”’s perception and reading of images, this “specialized” reading should be searched beyond the simple adjacency between the textual scene and its corresponding illustration.
3. The text’s main *topoi* or topics could be paralleled with the most frequently annotated *visual foci*.
4. The tradition of placing *the caption/* the title under the image (presented by Julia Thomas) might involve a change in perceiving the importance of elements in the image:
 - a. It might signal that the order of decoding illustrations is from content (image) to title, and not from title to content (as in the case of reading texts);
 - b. It might signal that the order of decoding illustrations - subsequently, the hierarchy of elements in the image - is given by the position of the title/ caption (bottom - up) and by the slow reading of the title/caption (left - right). It is not the meaning of words that is important in the caption, but the arrangement and the visuality of words. Such conjecture was suggested by a recent rereading of Thomas Mann’s *Doctor Faustus*, which brought up the issue of the musical sheet’s visuality: in a nutshell, the complexity of a musical composition could be grasped through mere visuality and not through the decoding of notes. Accordingly, the process of decoding the caption serves also as a basis for hierarchy, of classifying the importance of visual elements that are turned by the gaze into “foci” (F1, F2, F3, ...). By taking into consideration the normal position of the caption, the possible hierarchy of illustration zones: bottom left – bottom right – up left – up right.

Evaluation of NLP tools used (if relevant):

Not relevant. However, it is relevant to mention that, with the quick development of large language models, the platforms such as ChatGPT3 or ChatGPT4 will be the best solution to tailor projects and methods for literary corpora in lesser resourced languages.

Bibliography

Bode, Katherine. 2018. *A World of Fiction: Digital Collections and the Future of Literary History*, Ann Arbor: University of Michigan Press.

Kestemont, Mike & Luc Herman. 2019. “Can Machines Read (Literature)?”, in *Umanistica Digitale*, n.5, 2019, DOI: <https://doi.org/10.6092/issn.2532-8816/8511>.

Krieger, Murray. 2019 (1992). *Ekphrasis: The Illusion of the Natural Sign*, Baltimore: Johns Hopkins University Press.

Manovich, Lev. 2012. “How to Compare One Million Images?”, in *Understanding Digital Humanities* edited by David M. Berry, 249-279. London: Palgrave Macmillan.

McGann, Jerome. 2014. *A New Republic of Letters. Memory and Scholarship in the Age of Digital Reproduction*, Cambridge, Massachusetts – London, England: Harvard University Press.

McKenzie, D.F. 1999. *Bibliography and the Sociology of Texts*, Cambridge: Cambridge University Press.

Ning, Wang. 2009. “An “iconological turn” in literary and cultural studies and the reconstruction of visual culture”, *Semiotica* 176–1/4(2009), 29–46.

Panagiotidou, Maria-Eirini. 2020. *The Poetics of Ekphrasis. A Stylistic Approach*, London: Palgrave-Macmillan.

Wagner, Peter (ed.). 1996. *Icons-Texts-Iconotexts. Essays on Ekphrasis and Intermediality*, Berlin-New York: Walter de Gruyter.

Webb, Ruth. 2009. *Ekphrasis, imagination and persuasion in ancient rhetorical theory and practice*. Farnham: Ashgate.

Nørgaard, Nina. 2019. *Multimodal Stylistics of the Novel More Than Words*. New York and London: Routledge.

Lewis, Ann. 2009. *Sensibility, Reading and Illustration Spectacles and Signs in Graffigny, Marivaux and Rousseau*. London: Legenda. Modern Humanities Research Association and Routledge.

Brake, Laurel and Marysa Demoor. 2009. *The Lure of Illustration in the Nineteenth Century Picture and Press*. London: Palgrave Macmillan.

Male, Alan. 2017. *Illustration. A Theoretical and Contextual Perspective*. Second edition. London-Oxford-New York: Bloomsbury

Sullivan, Edmund J. 2016 (1921). *The Art of Illustration*. Mineola, New York: Dover Publications.

Mitchell, W.J.T. 2005. *What Do Pictures Want? The Lives and Loves of Images*. Chicago and London: University of Chicago Press.

Mitchell, W.J.T. 1986. *Iconology. Image, Text, Ideology*. Chicago: University of Chicago Press.

Piper, Andrew 2023, “What do characters do?”, in: *Conference Reader: 2nd Annual Conference of Computational Literary Studies (CCLS2023)*, edited by Christof Schöch, Élodie Ripoll, Evelyn Gius and Peer Trilcke. JCLS / Zenodo. <https://doi.org/10.5281/zenodo.8093598>.

Guhr, Svenja and Mark Algee-Hewitt 2023, “What’s that Scary Sound?”, in: *Conference Reader: 2nd Annual Conference of Computational Literary Studies (CCLS2023)*, edited by Christof Schöch, Élodie Ripoll, Evelyn Gius and Peer Trilcke. JCLS / Zenodo. <https://doi.org/10.5281/zenodo.8093598>.

Thomas, Julia. 2017. *Nineteenth-Century Illustration and the Digital: Studies in Word and Image*, London: Palgrave-Macmillan.

Oprescu, George. 1945. *Grafica românească în secolul al XIX-lea*. Volumul 1. Bucharest: Fundația regala pentru literatura și arta.

Bădicioiu, Laurențiu. 2021. *Repere ale comicului românesc în revistele satirico-umoristice din perioada interbelică*. Bucharest: Bucharest University Press.

Genette, Gérard. 1997 (1982). *Palimpsests: Literature in the Second Degree*. Trans. by C. Newman and C. U of Nebraska Press.

Ardila, Alfredo and others. 2010. “Illiteracy: The Neuropsychology of Cognition Without Reading”, *Archives of Clinical Neuropsychology*, Volume 25, Issue 8, December Pages 689–712, <https://doi.org/10.1093/arclin/acq079>.

<https://openai.com/dall-e-2/>

<https://dallery.gallery/dall-e-ai-guide-faq/>.

<https://poplite.org/>.

<https://proiectulbrancusihairo.wordpress.com/>.

<https://www.dracor.org/>.

<https://distantreading.github.io/ELTeC/>.

<https://stablediffusionweb.com/#demo>.

<https://www.superannotate.com/>.

To be sent to marco.raciti@dariah.eu

