



Visualization of Categorization: How to See the Wood and the Trees

Ophir Münz-Manor <ophirmm_at_openu_dot_ac_dot_il>, The Open University of Israel  <https://orcid.org/0000-0001-6333-345X>

Itay Marienberg-Milikowsky <itay_dot_marienberg_at_gmail_dot_com>, Ben-Gurion University of the Negev  <https://orcid.org/0000-0002-1150-7259>

Abstract

In the article, we present, theorize and contextualize an investigation of figurative language in a corpus of Hebrew liturgical poetry from late antiquity, from both a manual and a computational point of view. The study touches upon questions of distribution and patterns of usage of figures of speech as well as their literary-historical meanings. Focusing on figures of speech such as metaphors and similes, the corpus was first annotated manually with markers on papers, and a few years later it was annotated manually again, this time in a computer-assisted way, following a strictly categorized approach, using CATMA (an online literary annotation tool). The data was then transferred into ViS-À-ViS (an online visualization tool, developed by Münz-Manor and his team) that enables scholars to “see the wood” via various visualizations that single out, inter alia, repetitive patterns either at the level of the text or the annotations. The tool also enables one to visualize aggregated results concerning more than one text, allowing one to “zoom out” and see the “forest aspect” of the entire corpus or parts thereof. Interestingly, after visualizing the material in this way, it often turns out that the categories themselves need to be re-assessed. In other words, the categorization and visualization in themselves create a sort of hermeneutical circle in which both parts influence one another reciprocally.

Through the case study, we seek to demonstrate that, by using correct methods and tools (not only ViS-À-ViS but others also), one can ultimately use visualization of categorization as the basis for what might be called *established speculation*, or *not-trivial generalization*, which means, an interpretative act that tries to be based on clear findings, while at the same time enjoying the advantages of “over interpretation”. This approach, we argue, enables one to see the trees without losing sight of the wood, and vice versa; or “to give definition” – at least tentatively – “to the microcosms and macrocosms which describe the world around us” [Weingart 2022], be they factual or fictional.

Four talked about the pine tree. One defined it by genus, species, and variety. One assessed its disadvantages for the lumber industry. One quoted poems about pine trees in many languages. One took root, stretched out branches, and rustled.^[1]

Dan Pagis

Introduction: Figurative Language in Late Antiquity Hebrew Liturgical Poetry – between the Manual and the Computerized

The experiment described and discussed in this article began with Münz-Manor’s doctoral dissertation entitled “Studies in the Figurative Language of Pre-Classical Piyut”, submitted to the senate of the Hebrew University of Jerusalem in

2006 [Münz-Manor 2006a]. The dissertation explored the use of metaphor, simile, epithets, and other figurative devices in the earliest stratum of Hebrew liturgical poetry, composed in the Galilee in the fifth century of the Common Era, also known as *piyyut*, a Greek loan word meaning poetry. The corpus includes all the known poems from that time period and was recovered from medieval manuscripts of the Cairo Genizah. It includes 211 poems by various poets, most of them anonymous, with a word count of 47,556. Earlier studies suggested that, unlike in many other corpora of poetry, figurative language in early piyyut is rather marginal, and the main goal of the dissertation was to corroborate (or refute) this rather intuitive assertion and to some extent to quantify it. At the same time, the dissertation had a qualitative and literary goal, namely to carefully analyze the poems by reading them closely, and to draw poetic conclusions from the usage, or lack thereof, of figurative language by the poets [Münz-Manor 2011].

At the beginning of the century, long before the computational approach began to bring its influence to bear on the Israeli academy, the method used was rather simple.^[2] The poems in the corpus were read, word by word, using colored highlighters to mark the different figurative devices, such as metaphors, metonymies, and similes. This was painstaking, to be sure, but by the end of a process that took several months, the entire corpus was annotated. In hindsight, although there existed a notion of a clear-cut categorization of figurative language in the corpus, no well-defined *tagset* was created. To put it another way, even though this phase was managed in a relatively systematic way, it was not based on ontological thinking in the full sense of the term. The major figurative devices were annotated in an entirely flat hierarchy of metaphor, metonymy, synecdoche and simile. The use of these concepts was of course based on a strong theoretical foundation, but the definitions bore some flexibility. Indeed, at the time, the notion of a *tagset* as a defined and more or less stable logical representation of the categories was not an idea Münz-Manor had thought of.

As previously noted, in the second stage, an analysis was carried out in which different examples were drawn from the annotated corpus. Here, a more detailed analysis of literary phenomena was carried out, such as metaphors with biblical connotations [Münz-Manor 2006b], similes that were created due to structural features of the poems [Münz-Manor 2009], figurative epithets and the like [Münz-Manor 2016]. In other words, only after annotating the entire corpus on a very rudimentary level was the *tagset* enhanced, in a process which was neither efficient, systematic nor clever. At any rate, the main phenomena that were discovered were grouped into chapters and analyzed in depth. Despite this there remained a gap between the scope of the textual infrastructure prepared for the research and the doctoral chapters, which were theoretical and interpretive in nature, and did not necessarily exhaust the quantitative potential that was encapsulated in the preparatory work.

The conclusion of the dissertation was that the use of figurative language was indeed relatively marginal in the corpus, but that at the same time the usage that did occur was interesting and innovative. Although most of the discussion and conclusion was qualitative, a short passage did touch upon the quantitative aspects, even if in a scant way:

Out of approximately 5000 poetic lines, only 1000 contained figurative devices of any description, namely 20 percent. Moreover, the number of lines that contained developed figurative usage, which in most cases means a usage that runs across the entire line, was scarce, only 100.... the piyyutim [= piyyut in the plural] employ almost exclusively metaphors; the number of similes is surprisingly small, as well as the number of metonymies and synecdoches [Münz-Manor 2006a, pp. 238–239].

These comments were not further developed or expanded upon, and remain almost the only reference to the quantitative aspects that emerged from the painstaking preparatory work. In retrospect, this fact should probably be understood as a reflection of the well-known hesitancy of literature studies with regard to numbers [English 2010, p. xiii] [Goldstone and Underwood 2014, 359–362]. Many of the dissertation's chapters were published as journal articles and it seemed that this scholarly chapter in the author's career had ended.

That being said, fifteen years later, as our shared interest in computational literary studies grew, we came across CATMA (Computer Assisted Text Markup and Analysis), a web annotation tool developed at the University of Hamburg and now at the Technical University of Darmstadt [Gius et al. 2022], and a new scholarly project emerged. CATMA enables the annotation of every element in the text (and not the text as a whole – an issue that we will discuss later) according to tagsets defined by the user. Once the tagging has been accomplished, the user can execute queries

concerning both the text and the tagged text. Working with this tool, the first thing that caught Münz-Manor's eye was that the interface seemed almost identical to his dissertation's print-outs, namely the text highlighted with colors to depict the different annotations. But, unlike the initial work which was entirely manual, the main advantage of CATMA, or of any other computer-based annotation tool, is the ability to run basic or complex queries and to retrieve data. Thus, for example, in a hypothetical digital project equivalent to the doctoral thesis described earlier, the researcher could retrieve in seconds, and for the first time, a distribution of the figurative devices in the entire corpus, whether in a set of poems or a specific one; a list of frequent metaphoric words; or a percentage of figurative versus literal words. Such data is foundational for a combined qualitative-quantitative research. Moreover, the un-dogmatic conceptual principle of CATMA [Horstmann 2020] enables the annotators to enhance, refine or redefine the tagset on the go.^[3] As previously noted, the manual tagging during the initial phase of writing the dissertation forced Münz-Manor to use a one-dimensional and rigid *tagset* that, on the one hand, facilitated the tedious endeavor but, on the other, restricted the potential for complex analysis.

We began to wonder what would happen if CATMA were used to explore figurative language in *piyyut*, transforming the hypothetical idea described above into a reality. This experiment would have twofold importance; firstly, it would provide another perspective on the literary question at hand, and secondly it would serve as an unusual methodological trial of digital versus analogue approaches. We decided to run an experiment: we would “upload” the manual annotations from the doctoral print-outs to the system and see where it would take us [Münz-Manor 2022]. Here we took another important decision, which was to stick strictly to the original annotations in order for the experiment to be as sound as possible. As literary scholars, who cannot ignore the subjective nature of interpretation, it was an intuitive act: the decision whether a word is figurative or not is a hermeneutical one, and different scholars would tend to annotate the word differently. Moreover, even the same annotator might interpret the word differently at a later date or in a different context. It was because of this that we decided not to touch the original annotations when migrating from the analogue dissertation to the digital experiment; but we shall return to this point towards the end of this article.

The results of the experiment will be presented in two sections. In what follows we discuss some general quantitative results, then we continue with a more theoretical discussion of the annotating double experience and its function in the research development. Finally, we present our visualization tool and follow with another case study that builds upon the capabilities of the tool.

Case Study (A): Corroboration and Refutation

Once the corpus of poetry and the annotations were uploaded to CATMA it was very easy to run basic queries and to (re)assess some of the key findings of the dissertation [Münz-Manor 2022, 370–373]. The most pressing question was whether figurative language in the corpus is marginal or not. Interestingly, it turns out that the question is not that simple. Let us consider some numbers: the corpus includes 47,556 words, out of which 5113 are figurative. In other words, the share of figurative words in the corpus is 11%, a small fraction to be sure. In that sense, we can say that the quantitative results corroborate the qualitative ones from the dissertation. But “small” in comparison to what? In the dissertation this question could not have been answered and probably did not seem worthwhile at the time. In order to correlate the results, we examined them against two new corpora that have now been annotated for figurative language: the Book of Psalms from the Bible and the poetry of Pinchas the Priest, a ninth-century poet from Tiberias. These corpora were selected for a particular reason; the Book of Psalms is the foundational corpus of Hebrew poetry and *piyyut* has a direct connection to it even if the late antique poets sought to distance themselves from the canonical text. The oeuvre of Pinchas, on the other hand, builds heavily on the early *piyyut* and represents the inner development of this specific literary tradition after three hundred years of development. Simply put, the original corpus of the PhD dissertation was placed in diachronic comparison with an early and a late corpora.

The percentage of figurative words in the Book of Psalms is 14% and in the poetry of Pinchas 9%. So, on comparing the two, it appears that the original corpus of the dissertation was not so under-figurative as seemed at first sight. In fact, it stands in between the classical book of Hebrew poetry, the Psalms, and a later poet that belonged directly to the literary tradition of *piyyut*. It is hard to draw a conclusion solely from these quantitative observations with regard to the history of Hebrew poetry in its diachronic development as far as the usage of figurative language is concerned, yet it definitely

marks a promising scholarly path for future studies.

In other cases, the digital reexamination refuted findings in the dissertation. One of the meaningful categories for the analysis in the dissertation was the relation between figurative language and the specific genre or sub-genre of the poems. Piyut is characterized by a strict generic classification and naturally some (sub)genres tend to use figurative language more than others. The quantification of the results allowed one to examine easily what the figurative density (the ratio between figurative and non-figurative words) of each genre might be, and here, too, the results corroborated some of the initial findings while refuting others. Most notably is the genre of the Hosha'not (hosanna), which ranked at 25% of figurative density. That is to say, in general, every fourth word is figurative. If we compare this phenomenon to the overall figurative density of the corpus, which is 11%, it really stands out. Curiously, in Münz-Manor's dissertation it was noted, in passing to be honest, that this genre is low on figurative language [Münz-Manor 2006a, p. 61]. Where did such an incorrect assertion come from? The Hosha'not are very short and in the style of a litany, a list of attributes of the people of Israel. So even if many of these epithets are figurative, the poem as a whole does not sound very figurative. Here we encounter an interesting case where the quantitative and the qualitative represent separate lines that do not meet while both are valid. Quantitatively, there are many figurative words, however, their aesthetic impression is low and possibly does not call for a detailed literary analysis; or does it? We do not think we can or should decide on such matters, but it demonstrates quite nicely that opening the door to quantitative considerations expands the literary toolkit and the hermeneutical possibilities.

10

Manual Annotating in the Digital Age as a Close Reading Practice

At first glance, the transition described above from paper-based manual annotating to computer-assisted manual annotating, which undoubtedly provides better analytical capabilities [Jacke 2018], should lead to research which can be measured quantitatively, and whose conclusions will – almost by definition – be more robust. Simply put, although the research procedure was essentially the same, what Münz-Manor was able to do for the first time with the digital tool allowed him to say clearer things about the entire corpus. In other words, even in this seemingly modest case, the transition to a digital tool might well have coincided with a certain and widespread perception of the adopting of the computational option as a move in the direction of objectivity, scientification, and the like. However, as we shall see, and as other scholars have already argued, the picture is more complex, since, even in such a case, this move has a price [Rieder and Röhle 2012].

11

To begin with, we must ask ourselves if, and to what extent, the computer-assisted manual annotating is really so conceptually similar to the non-computerized manual one. Manual annotation in the digital humanities is sometimes described as an area of the good old close reading within the new realm of distant reading. This is both true and not true: it is true because, unlike some algorithmic analytic methods which are now very common in computational literary studies, human digital annotation is based on sensitive attention to every detail in the text, and is definitely a way of reading. And it is not true, because, unlike close reading, annotation in digital humanities is usually supposed to be based not only on clearly defined categories, but also on much more developed systematic thought, which, in turn, affects the scope of annotation: Often, a modest and well-defined annotating task is expanded to encompass many more phenomena, in an attempt to fully exploit the opportunities inherent in more and more annotated data. Traditional close reading is not adversely affected by this consideration. This problem, in brief, stems from the multifaceted nature of digital modeling operations – including annotating, as a form of modeling: such operations sometimes seek to focus on a specific pre-defined research question, at the same time as seeking to prepare the research object for future analysis, based on as-yet-unknown questions [Flanders and Jannidis 2016]. As a result, manual-annotation-based projects tend to lack (or, better, to challenge) another important and characteristic feature of close reading: sampling. As Paul Fleming puts it, “an essential element of close reading relies not just on the quality of the reading performed, but also on the example chosen. It has to be the right example” [Fleming 2017, p. 437]. It is not surprising, then, that in Münz-Manor's initial work, a sort of disconnect between the annotation process and the body of the dissertation chapters becomes apparent: while each chapter chooses the “right example”, these examples weren't chosen for their representativity, but because they were interesting or important to the author for some reason or other. The aim of the annotating process, by contrast, was only to prepare an infrastructure, and in the end that infrastructure remained largely in the shadows.

12

How much room is there for sample-based research in a computationally-based annotation process? Could a particular concept of categorization support it? Sampling of a certain portion of a larger corpus is certainly justified in the context of traditional literature research. It reflects the gap between the amount of data prepared and the relatively free discussions that do not necessarily derive from it. It is, among other things, what gives this kind of research freedom of interpretation and room for creative conceptualization which is not necessarily less reliable. Choosing the right example, then, is one of the most effective ways of moving between detail-orientedness and generalization, which is the traditional way to see the wood for the trees. However, a quantitative approach that takes itself seriously may necessarily limit itself from using such a method, even in relatively simple cases which do not give up on human reading in favor of completely computational automation. Instead of highlighting the most interesting cases, it will highlight only those that “actually” – that is, statistically – represent something broader than themselves. That being said, how can one bridge this gap in computational literary studies, without giving up systematic category-based annotation? Even when annotation is driven by interpretive considerations, at the end of the process a layer of information is obtained that is perceived as rigid, whose reprocessing with statistical tools may not reveal its interpretive richness. Doesn't the meticulousness of manual annotating avoid the additional interpretive step beyond what emerges directly from the accumulation of data? Do generalizations, in the digital age, have to be more modest, more data-driven?

13

To put it another way, is the transition between the carefully annotated data to what can (and should) be said about such data still possible? This becomes even more complicated when – typically of digital humanities – “the wood” is not one text, but many, which are now being observed and analyzed through a well-developed abstract category system.

14

Taking the analogy further, if the text is a tree, then, when we label its parts, we also give names and definitions to the shape of its branches, to the texture of its leaves, to the taste of its fruits. But since our corpus is a forest that contains many texts – many “trees” – as is usually the case in computational literary studies, when we then want to understand the forest as such, we can no longer pay attention to every branch, leaf, and fruit. Generalizing about a tree means talking about the entirety of its branches, leaves, roots, and fruits. Generalizing about a forest means talking about the entirety of its trees. At this point, the connection between the parts of individual trees and the entire forest might become increasingly blurred. To be clear, this is not another version of Franco Moretti's distant reading, because, unlike Moretti, when we seriously annotate text after text we *do* interpret it. However, in the end, even if we find the most creative ways to build a database of texts that classifies them on the basis of a complicated calculation of all their annotations and internal relationships, in the end, some degree of reduction is almost inevitable.

15

Admittedly, it is complicated even at the level of one tree taken as a whole: Anyone who has experienced the transition from text annotation to text classification knows that it is not an easy transition. Text annotation is a device designed to break down texts; text classification – and genre classification is a good example of this – is a device designed to connect their parts together, in order to achieve at least one generalization regarding the text – one that facilitates comparing it to other texts on a well-defined basis.^[4]

16

If we follow Andrew Piper's attitude to generalizations, as described in “Can We Be Wrong? The Problem of Textual Evidence in a Time of Data” (2020), it seems that the answers to the challenge phrased above would be relatively clear. In response to one of his book's manuscript's readers, who wrote that “this sounds awfully boring”, Piper playfully agrees: “It is! Generalizations should be boring, cautious, incremental, and slow” [Piper 2020, p. 60]. It is impossible to deny the logic of this approach, which expresses a central element in the research experience in the age of data. It is hard, however, not to feel the stark contrast between this approach and that of, say, Jonathan Culler, who offers the opposite perspective of the pre-computational age:

17

Moderate interpretation [...] though it may have value in some circumstances, is of little interest [...]. Many “extreme” interpretations, like many moderate interpretations, will no doubt have little impact, because they are judged unpersuasive or redundant or irrelevant or boring, but if they are extreme, they have a better chance [...] of bringing to light connections or implications not previously noticed or reflected on than if they strive to remain “sound” or moderate [Culler 1992, 110].

The question then is – can we do both? Is it possible to be content with modest generalizations, as Piper recommends, while at the same time going far into the bold realms of over-interpretations, as Culler recommends?

In answering this question, we assume that Culler is not talking about *false* generalization, but speculative ones: Thought-provoking generalizations that are not derived directly from the data, and whose relation to the data may even be questionable, but nevertheless reveal a different kind of truth. When we put the debate this way, with all due respect to data-driven research, there is no reason to deny the importance and the fruitfulness of more radical hypothesis-driven research – especially not in computational literary studies (or other sub-fields of digital humanities). And here, we believe, categories, which sometimes appear to serve as a means to enrich data in a relatively robust manner, do not have to be thought of as a hindrance – a device that prevents speculation – but rather as something to work with; as a valuable source of inspiration. Indeed, in what follows we introduce a tool that is designed, among other things, to enable just that: at one and the same time to take advantage of categorization while also fostering speculation.

18

Visualization as Hermeneutical Tool: Introducing ViS-À-ViS^[5]

One of the problems with annotating projects of the type that Wendell Piez described as “hermeneutic markup” [Piez 2010], and which CATMA reflects [Meister 2014] [Horstmann 2020], is that this type of tagging strongly encourages a focusing on details – on the parts of the tree – and often makes it difficult to see the whole picture, the “wood”. Even when the user analyzes the annotated material, their analysis is usually question-dependent, revealing a small subset of the dimensions of the tagging. At this point, we believe, the need for an advanced visualization of the hermeneutic metadata, that is, the manual annotations, becomes urgent – and this is where the tool that will be described below comes into play. While visualization can be question-dependent, it can also be used more freely. And despite all the problems associated with it, we assume, like many others, that this is one of the most important paths for distant reading; there is no other way to see all the details from above [Drucker 2011] . Recently Franco Moretti commented that no way has yet been found for an adequate synthesis of text-interpretation with corpus-visualization [Moretti 2020]. We hope that the idea presented here advances us towards a possible solution.

19

It should be noted in this context that, unlike close reading, where the scholar actually reads the text closely, distant reading – regardless of its specific history [Ross 2014] [Herrnstein-Smith 2016] [Underwood 2017] – is better understood as a metaphor. In other words, the reader extracts meaning from the corpus without actually reading it, à la Moretti. Analyzing text via visualization, then, is a perfect example of “reading” a text, without reading it at all. But, in our case, the fact that the visualization of the entire corpus is based on actual reading – on quite sensitive reading – makes this act of reading-through-visualization *a bit more* a way of reading in the original sense: the “trees”, if you like, are still there.

20

CATMA is a powerful and flexible tool that focuses on annotation, hence its visualization capabilities are basic. This reality was the incentive for the creation of a separate tool, ViS-À-ViS, that enables the user to visualize the annotated text, the annotations (by one or more annotators) and even the tagset itself [Münz-Manor et al 2020]. That being said, following Johanna Drucker’s call for more humanities-oriented visualizations [Drucker 2011], the tool seeks to provide users with more than mere visualizations, namely, it provides an interface that supports what might be called speculative hermeneutics, by allowing for a smooth transition between the smallest detail and the entire picture, and doing so in a playful manner. By this, we mean that the tool provides the opportunity to “play” with the visualizations of the individual texts, and group them in various configurations according to their similarity or dissimilarity, with the aim of providing the users with a fresh insight into the text and its annotations. One can think in this context of a partial similarity between this tool and the famous visualization tool Voyant [Sinclair and Rockwell 2016]: while Voyant focuses on the text, in a way that encourages free speculations, our tool does the same with annotations as the object of analysis.

21

In more than one way, ViS-À-ViS is following in the footsteps of another project in the CATMA ecosystem, Stereoscope.

^[6] According to the developers: “Stereoscope is a web-based software prototype for visualizing two core processes of literary studies: hermeneutic exploration of textual meaning and construction of arguments about texts” [Kleymann and Strange 2021]. According to Kleymann and Stange, the developers of the tool, its core principle is what they call

22

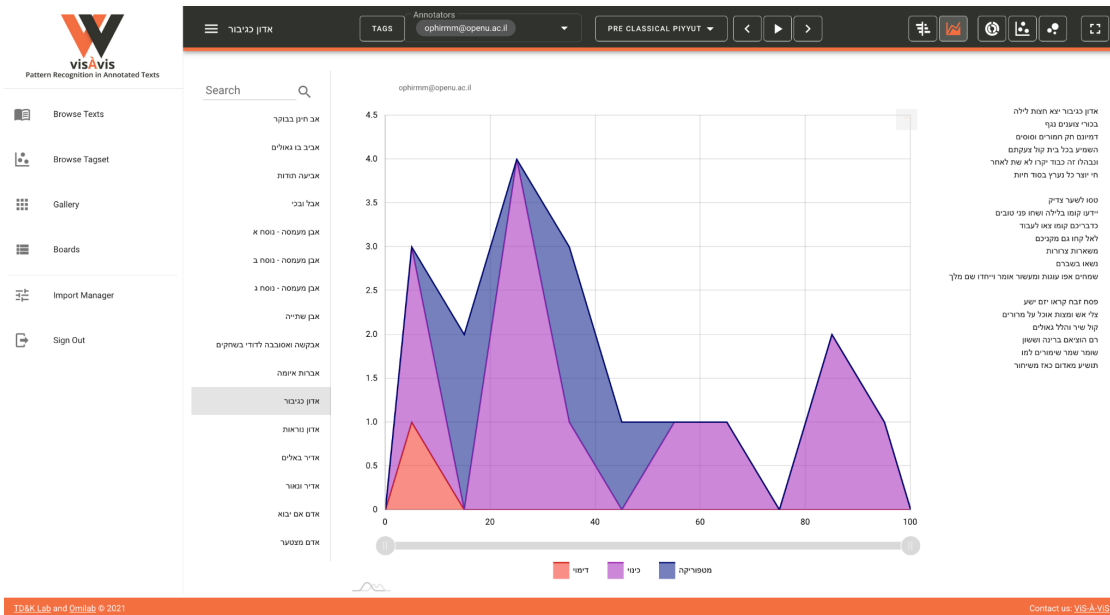


Figure 2. Stacked Area View

This visualization is similar on the one hand to the Gantt as it presents a “diachronic” (x-axis) overview of the annotations in the text. On the other hand, it provides a more nuanced representation of the dynamics between the various annotations on the “synchronic” (y-axis) level. Hence it enables the scholar to detect general patterns throughout the course of the text as well as patterns in the interaction between the tags at specific parts thereof. What is more, this visualization is interactive, and with a single click a tag can be added or withdrawn from the screen, thus enabling the user to play with the visualization and consider different views and interpretation modes. Here, one of the core features of the tool comes into play, one that gives it its experimental nature, namely, the ability not only to visualize the annotations statically but also to add or eliminate some of them in order to reveal different aspects of the texts and the patterns embedded within them.

26

Other charts offer the opportunity to present aggregative views. For example, in the sunburst chart, the user can visualize one text (Figure 3):

27

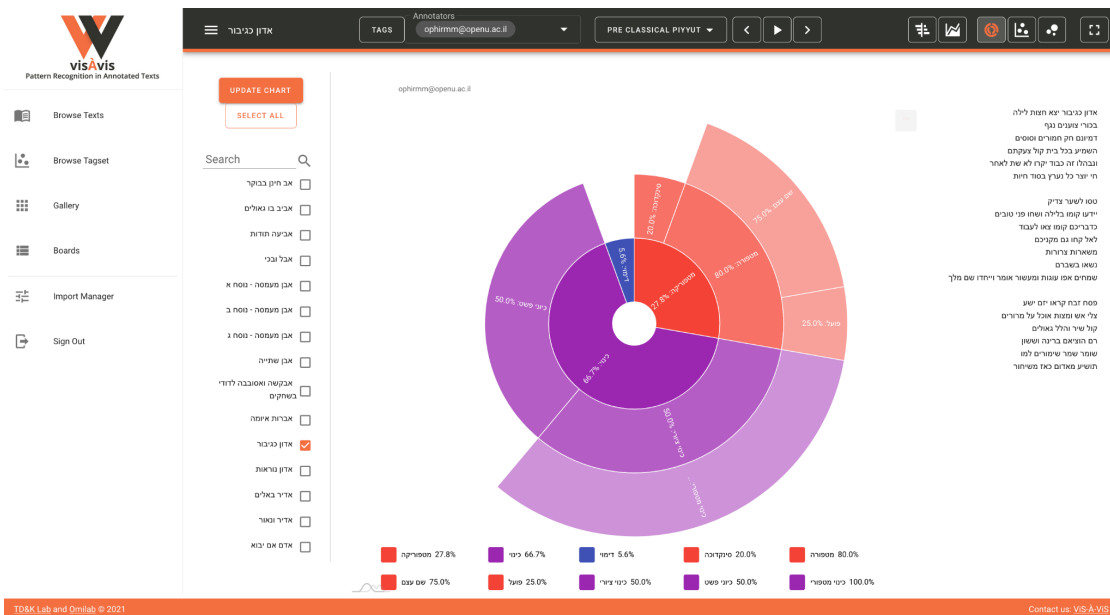


Figure 3. Sunburst View

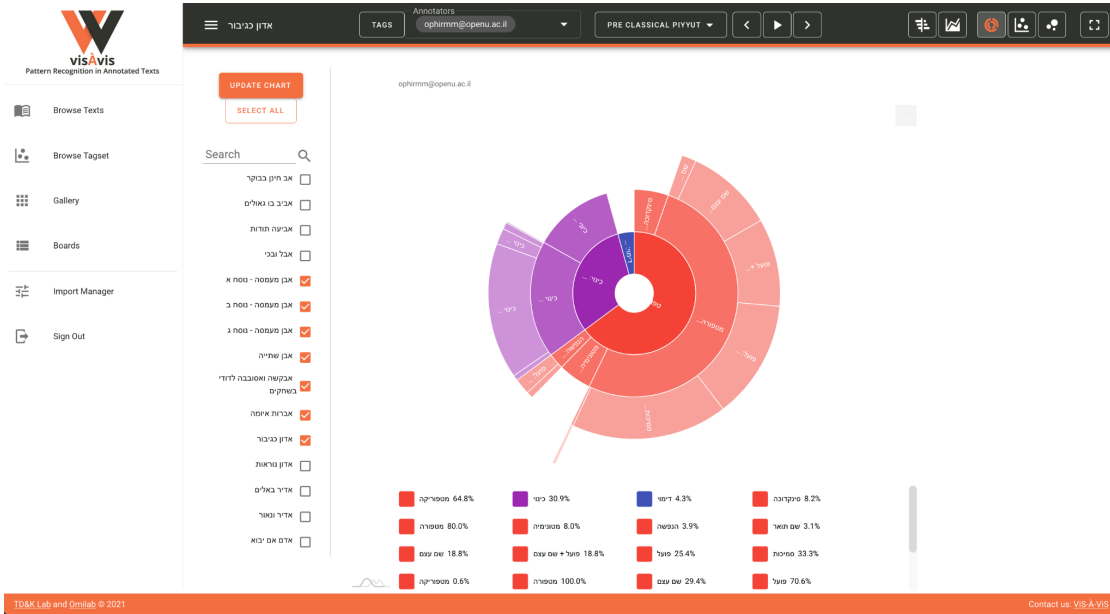


Figure 4. Sunburst View - Aggregated

In each case the interactive chart presents the distribution of the selected tags in terms of their percentage in the entire text(s) – a very useful feature that offers a quantitative appraisal of the annotations. In practice, this functionality provides a powerful analytical tool for the scholar, who can precisely equate or differentiate between different texts in the corpus or to compare different corpora. Here, too, the interactivity of the visualization offers the scholar a high degree of flexibility in the exploration process.

Browse Tagset is another meaningful visualization option and one that needs to be singled out in the context of this article. Large, hierarchical tagsets are hard to grasp and visualize in CATMA, and ViS-À-ViS offers an interactive representation of the tagset in a network-oriented chart (Figure 5):

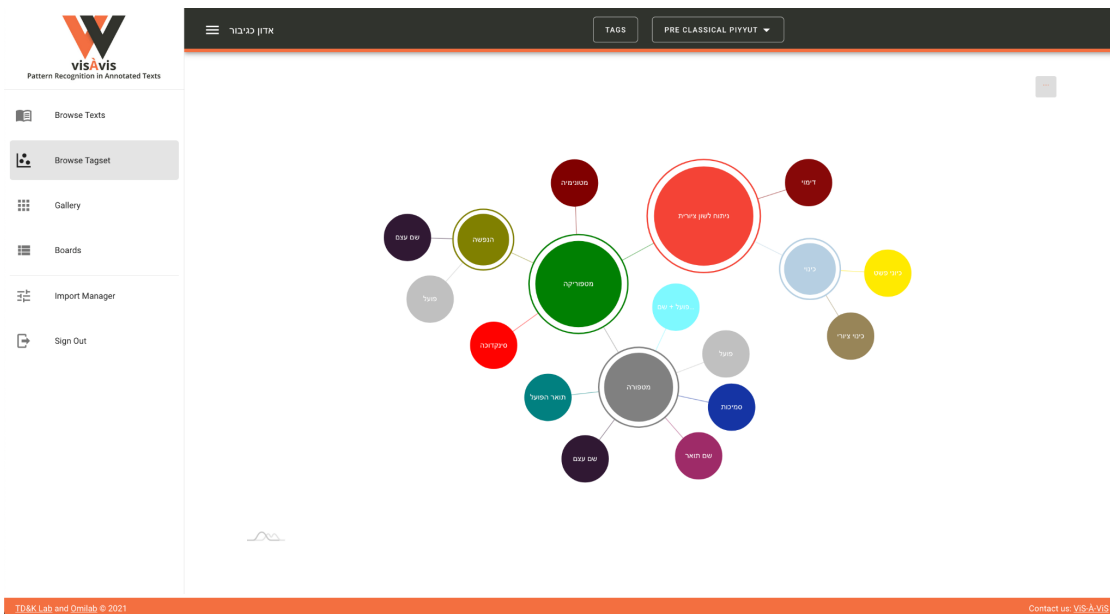


Figure 5. Browse Tagset View

Here a representation of the tagset is visible and the interactive chart enables one to collapse or uncollapse the various

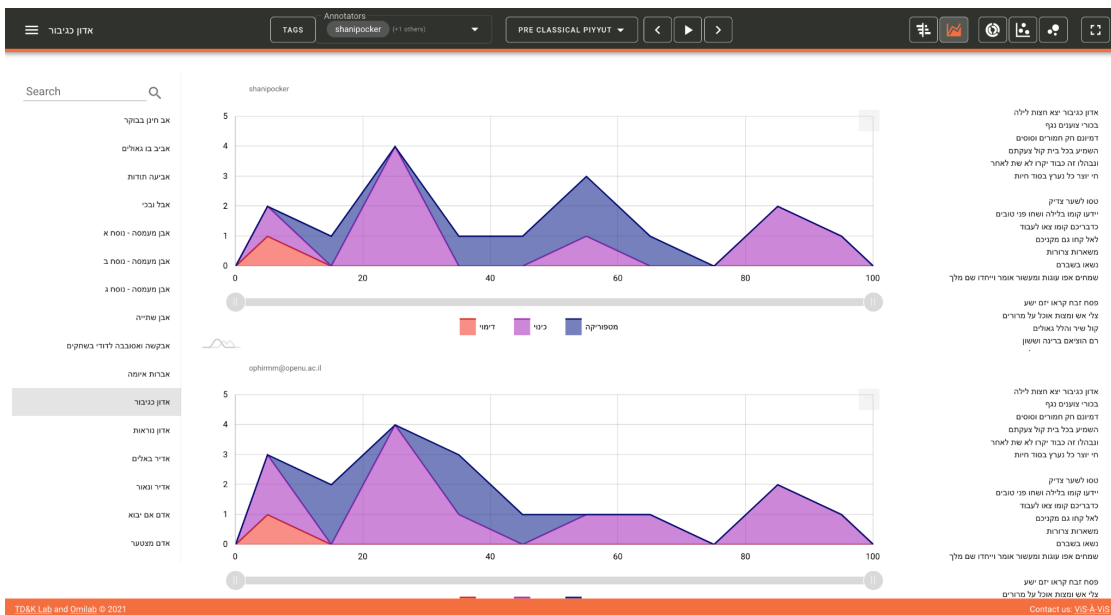


Figure 7. Multi-Annotator Mode

Inter-annotator agreement (or lack thereof) is a fundamental term in computational linguistics and natural language processing, fields that seek to minimize human variation and to create standards [Plank 2022]. In “traditional” literary studies variation is frequently the driving force of the hermeneutical process – and, as we argued elsewhere, we think it can and should be the same in the field of computational literary studies [Marienberg-Milikowsky et al. 2022]. Instead of collapsing the variation into a gold or a silver standard, ViS-À-ViS advocates for the visualization of multiplicity. In other words, instead of providing an answer concerning the “correct” annotation, it raises a question about it and promotes once more literary speculation. 34

ViS-À-ViS offers additional functionalities that enable the user to view miniature visualizations of different texts side by side (Figure 8): 35

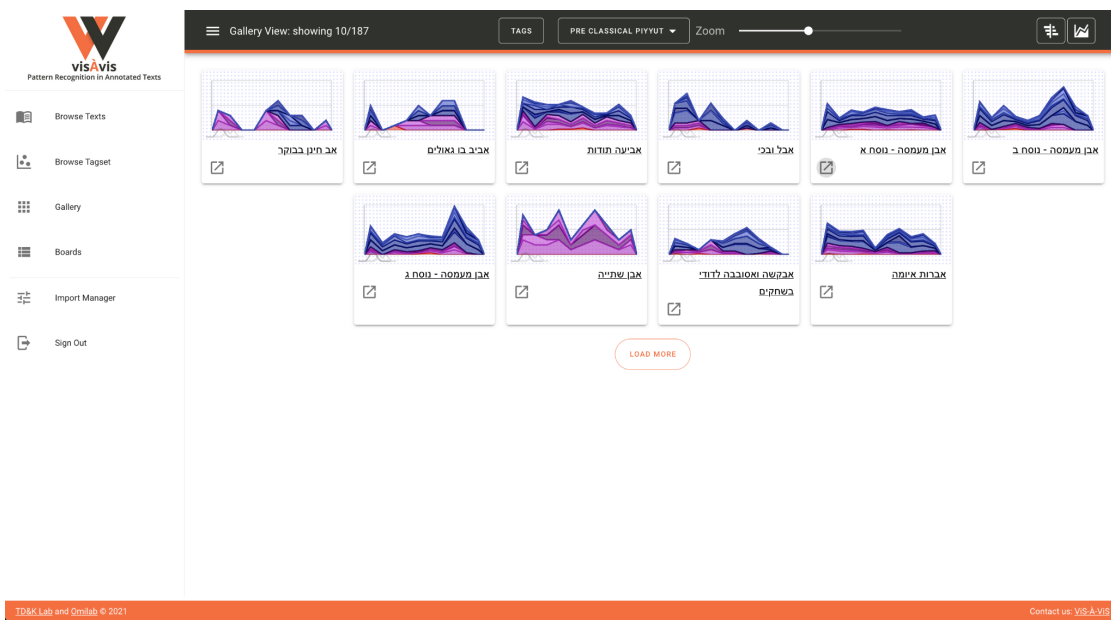


Figure 8. Gallery View

By aggregating the texts, the user can detect similar (or dissimilar) patterns and use the Boards option to collect and 36

divide them into different groups (Figure 9):

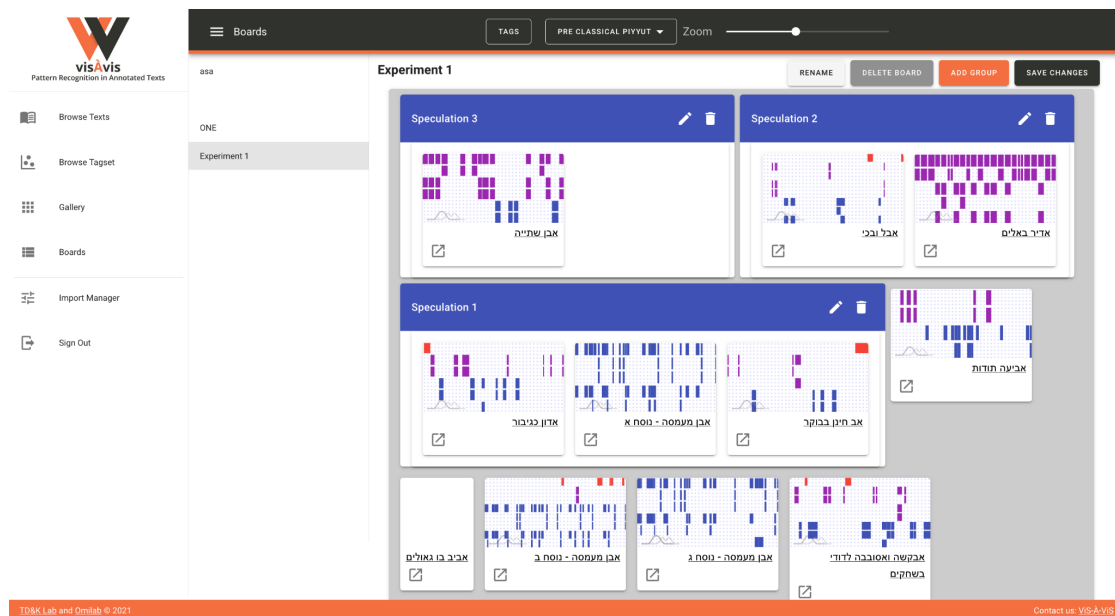


Figure 9. Boards View

This exemplifies perfectly the speculative potential of ViS-À-ViS that enables the user to group, re-group and de-group texts according to flexible categories defined by him- or herself. Indeed, in the next section, we will exemplify how using these features enabled us to detect an unusual pattern, to explore it, and reach an interesting scholarly conclusion.

37

Case Study (B): Hermeneutical Visualization

The process of playful visualizations can sometimes lead to new revelations. Using ViS-À-ViS' various modes the scholar can skim through the corpus and look for (recurring) patterns or other notable features of the tagged texts. Indeed, it is very difficult to frame exactly this exploratory phase, to predict how long will it take or to predict its fruitfulness. In this case we set out by using the Gallery View the Gantt and Stacked Area charts visualizations of the entire corpus, genre by genre. We then created some Boards with "interesting" visualizations of specific poems and compared them one to the other. The whole process took roughly two weeks and at the end, this cross-genre exploration highlighted an "unusual" similarity in one of the Boards and sent us back to one of the sub-genre of the corpus, the Seder Avodah (order of sacrifices). The genre is a lengthy poem for the Day of Atonement that narrates and associates the creation of the world and the history of the first generations with the sacrificial ritual of the high priest in the Jerusalem Temple [Swartz and Yahalom 2005]. The pattern that we recognized is characterized by a drop in the number of usages of figurative devices at a certain point in the poems, while towards the end of the composition the number rises again. The figures below visualize the annotations' distribution in four Avodah poems. The left-hand side represents the beginning of the text, and the right-hand side its end. The vertical axis represents the number of figurative usages, aggregated per ten-percent chunks of the texts (Figure 10, Figure 11, Figure 12, and Figure 13):

38

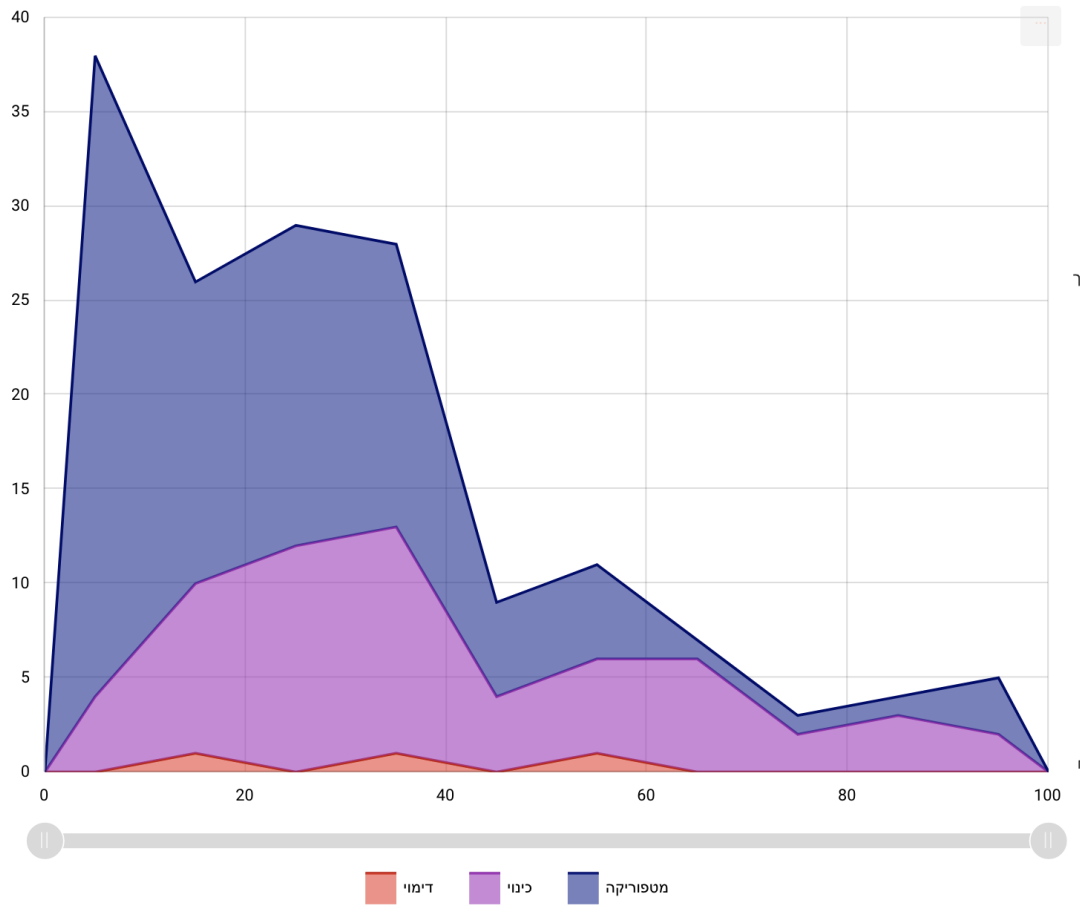


Figure 10. Distribution of figurative devices in "Ata konanta olam berov khesed"

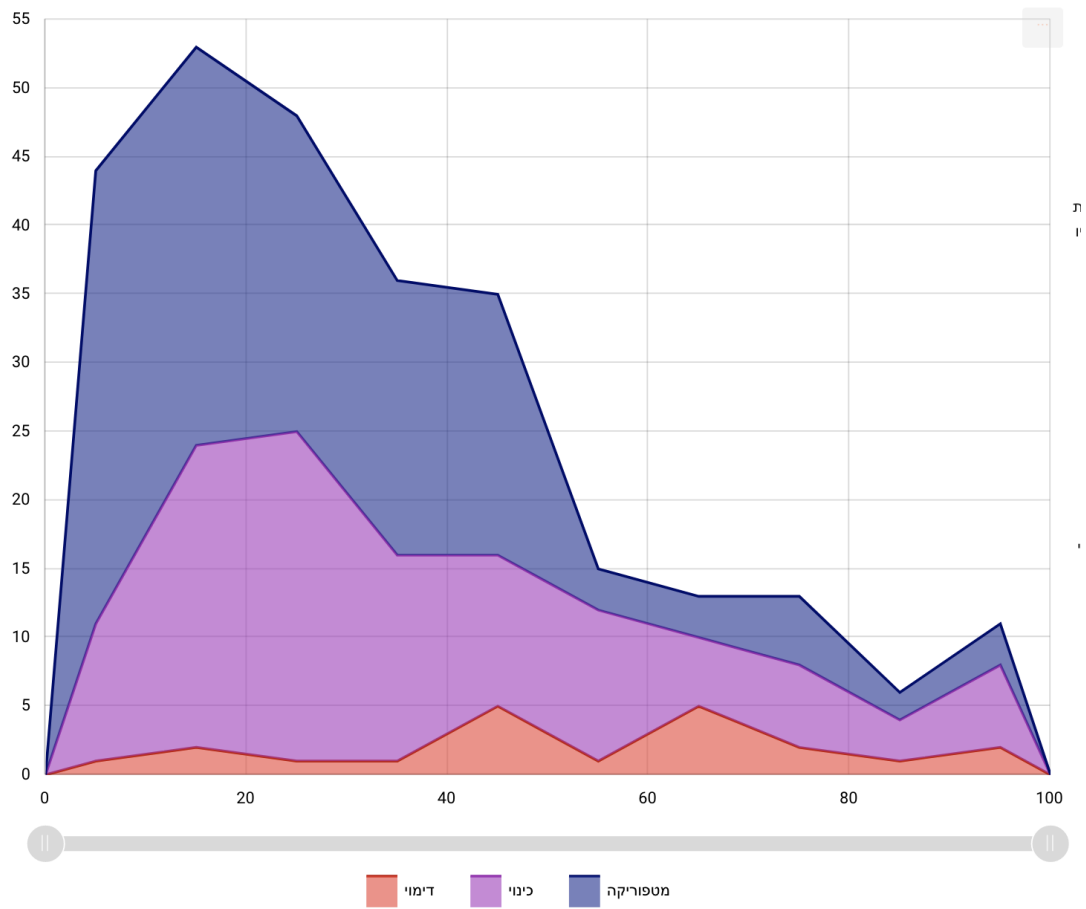


Figure 11. Distribution of figurative devices in "Azkir gevurot eloah"

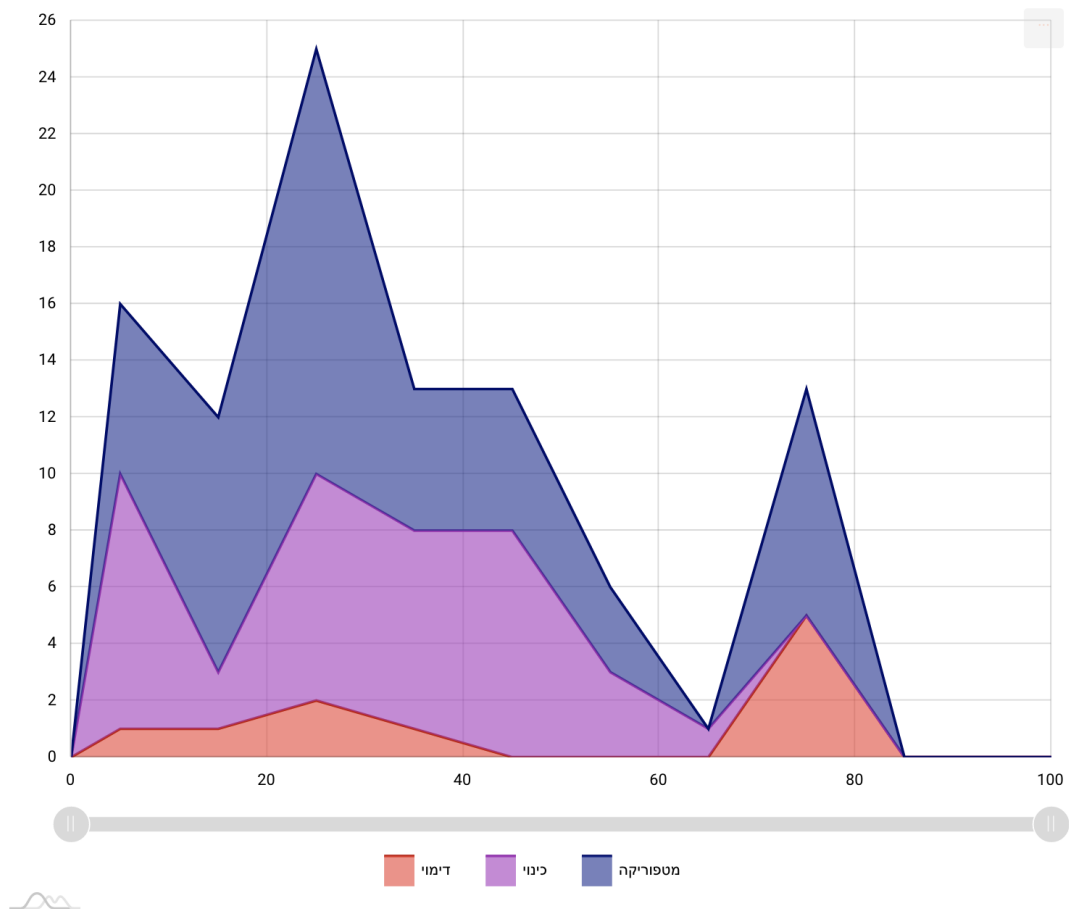


Figure 12. Distribution of figurative devices in "Aromem bekhol et"

Here, another important phase comes into play, in the interplay between the distant and close reading. We went back to the text and discovered that this shift appears in the transition from the description of the history of the world to the description of the cult of the high priest. This phenomenon had gone almost entirely unnoticed in scholarly discussion and was hinted only in passing in Münz-Manor's dissertation. Interestingly, in one poem this pattern breaks:

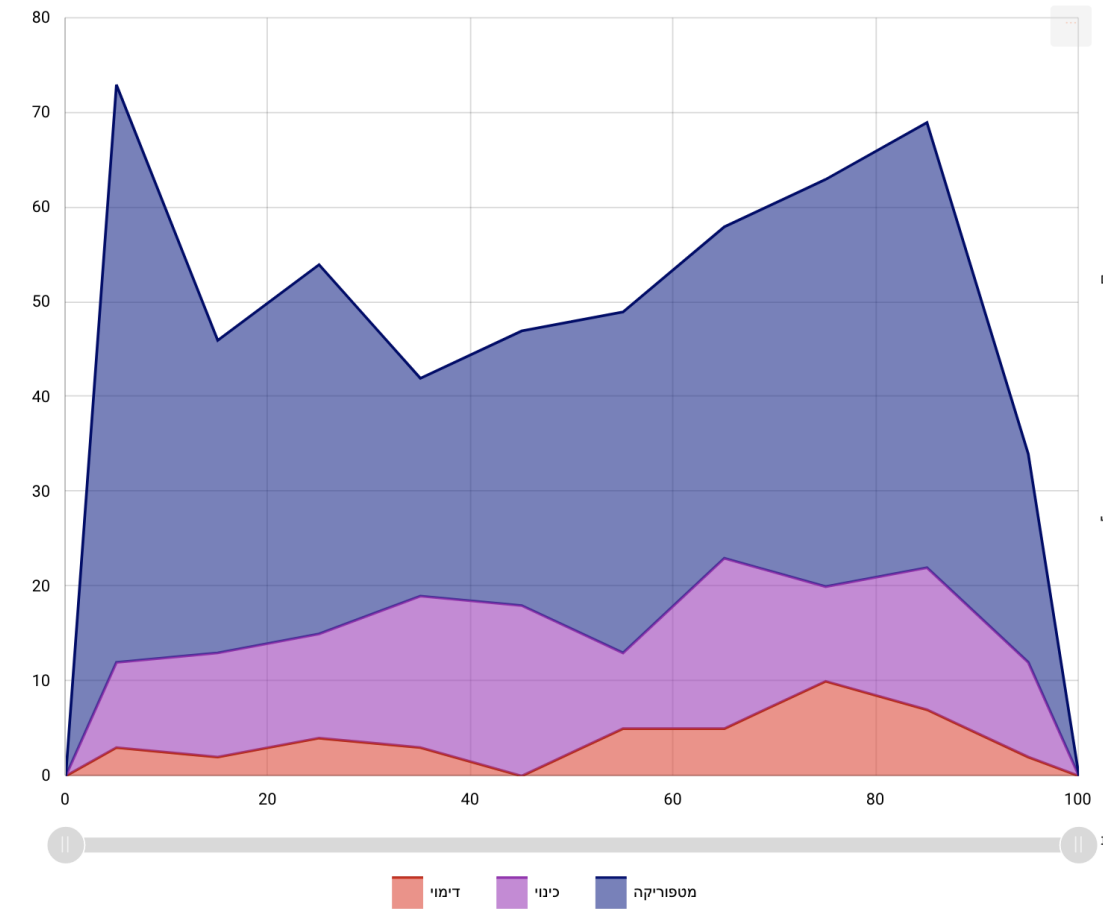


Figure 13. Distribution of figurative devices in "Az be'ain kol"

Interestingly, but not necessarily surprisingly. There is an ongoing dispute among scholars whether it is very early (ca. fifth century CE) or very late (ca. ninth century CE) and the structural analysis suggests that it is the latter [Münz-Manor 2022, 373–378]. 40

The case of the Avodah poems thus nicely illustrates the “zoom-in/zoom out” dynamic of close versus distant reading that such a tool offers. The discovery of the latent pattern began with a close reading of the poems and the manual annotation process; continued with a distant-reading visualization that exposed the patterns, without ignoring the fact that these patterns are based on a very sensitive reading; and concluded with going back to the text in order to interpret the finding – and sometimes reinterpret the text. This workflow demonstrates how the introduction of quantitative methods and data visualization, when done from a literary studies point of view, can expand and enrich the toolbox of the literary scholar. In future versions of ViS-À-ViS we envision an enhancement of the workflow by implementation of automated pattern detection, a task already underway [Münz-Manor et al 2020, 2–3]. Obviously, the automatic detection of patterns would not replace the scholar and the hermeneutical process, but it has the potential to broaden the interpretative possibilities and to make the workflow less cumbersome. 41

From the Analogue to the Digital: The Conclusions of the Reexamination (and a Note on Pedagogy)

Literary studies can benefit greatly from walking a moderate line, where one piece of evidence after another is gathered, and everything is done in a measured and balanced way, with openness to criticism, rather than indulging in generalizations based on rhetorical persuasion. Our approach here is not very different from Piper’s, although we believe that this argument goes quite naturally hand in hand with computational research projects that rely heavily on the systematic annotation of texts, rather than on machine learning algorithms, which – despite being based on pre- 42

existing annotated texts – emphasize the distance from the text itself. As long as we still wish to “read” – in a purely literal sense – this is probably the safest approach. It could also respond to Adam Hammond's justified call for the validation of research findings in the digital humanities [Hammond 2017]. However, the price of such a procedure is not only that too much focusing on details can hide the bigger picture, and sometimes even leave the researcher with the feeling that reaching it is impossible, or that it is insufficiently well founded, but also such research may lose some of its significance, its *raison d'être*, and so become, in one way or another, less meaningful. If we echo again what was presented before as a kind of (imagined) dispute between Piper and Culler, the question is how to find the golden path between a textual-anchored generalization on the one hand, and, on the other, a “healthy” speculation that will advance the research – with a calculated risk – into new and unexpected areas.

Here, we argue, a different attitude to categories can help: It is usually difficult to move directly from the smallest details to the bigger picture. We therefore suggest focusing on categories, as a research object located somewhere in the middle, and anyway requiring some abstraction – while bearing in mind that categories might have non-trivial relationships with both the details and the picture. Simply put, we propose a more playful approach to the epistemological status of categories, one that does not leave them solely as the cornerstone of the annotation process, but returns to examine them again and again from a critical perspective. From a more philosophical point of view, it can be said that this approach offers a back-and-forth movement between two different types of categorization: a categorization based on generalizations – that is, collecting and extracting all the data until categories that contain them correctly are defined – versus a categorization based on formalization, that is, by removing all actual data until an abstract category is formed (see: [Thomasson 2022, § 1.3]).

This approach also has a pedagogical aspect, which in our opinion is fundamental for the teaching of computational research to humanities students, and symptomatic of the subject at hand. We can demonstrate this from our shared experience in teaching manual digital annotating of literary texts. When we teach CATMA to our students – and we do it a lot – we are used to telling them that the main place where their traditional theoretical thinking is supposed to undergo computational operationalization is in setting the tagset, that is, in the translating of abstract concepts into a category system. The daring ones usually adopt the proposal and try to formulate a whole world of fuzzy ideas as a tagset, which, flexible as it may be, is more rigid by definition than the primordial thoughts with which the students set out. But then, when they are starting to annotate, they want to tag everything; they want to add new tags, or correct the old ones; and that almost always happens. In a way, they are good students. Once they are exposed to the power of systematic annotation, they become addicted to it. They fall in love with the straight-forward use of a category system: it offers them an attractive intellectual alternative in its solidity, in its precision, in its caution. They start like Culler, and end like Piper. When, later, they analyze their annotations through a concrete query, with or without some basic visualizations, they transform their annotations into something tangible, and the process comes to a head: metadata becomes data, and interpretation becomes a rigid text. “A Trojan horse”, is what Johanna Drucker (2011) once called visualizations with too much explanatory power. When we try to challenge the students and say – “now speculate, please!” – they find it difficult to achieve. They see, of course, the richness of the data; they see – hopefully – the preciseness and carefulness of the local annotations; after all the hard work they have invested, they can treat it as reliable information. This is all good. But the leap towards a certain kind of reduction, or more creative speculation, or a bit less grounded but nonetheless legitimate generalization, ends up being viewed with circumspection and as a serious challenge.

The tool presented here was constructed not only based on the lessons we had learned from moving the research discussed above from a paper-based environment to a computerized environment, but also with these students in mind. We argue that excessive textual anchoring of speculations and hypotheses comes with a price, and offer a visualization approach that tries to deal with this; not only in order to achieve the digital equivalent of “the right example” chosen – a speculative generalization of the entire “tree”, or “wood”, based on the careful annotation of its parts – but also in order to push the researcher to rethink or rephrase their categories, by moving as freely as possible between generalization-based categorization and formalization-based categorization. It might encourage the researcher to redefine his or her category system, and this time not only on the basis of detailed interpretation, but also on the basis of a (careful) reduction.

Finally, we want to offer a sort of thought experiment for a workflow that is not yet possible with the tool presented here,

but in our opinion can outline one of the next steps in research. What if we take to an extreme the promise of a tension between the data-driven attitude and hypothesis-driven attitude. Many human annotators hope that at some point the machine will learn how to annotate the text, and continue automatically with their own manual sensitive work; actually, some annotation tools already do this.^[8] But what if we try it the other way around, and speculate about the opposite possibility, where we start from the forest, from generalization about the trees or the wood as a whole – that is, from a hypothesis – and only then present it to the annotator, while asking him or her: Can you reject this generalization? Can you interpret the text differently so that the data will ultimately undermine the validity of the generalization?^[9] Or, alternatively, can you explain the validity of this claim by abstraction instead of generalization? Such an exercise's purpose is not to encourage the interpreter to manipulate (in a negative sense) the data, but to challenge him or her in a very concrete way; to aid him or her in thinking differently about the data and generalizations associated with them.

Trees and woods, after all, may look different at any given moment, and there is no reason to confuse systematicity with stability. It might be helpful, perhaps, to quote here Roland Barthes' adaptation of Nietzsche, who said that "a tree is a new thing at every instant: we affirm the form because we do not seize the subtlety of an absolute moment". And Barthes: "The text too is this tree whose (provisional) nomination we owe to the coarseness of our organs. We are scientific" – please note – "because we lack subtlety" [Barthes 1975].

47

Acknowledgment

This research was generously supported by grant No. 1223 from the Israeli Ministry of Science and Technology.

48

Notes

[1] Dan Pagis, Conversation. Translated from Hebrew by Stephen Mitchell; <https://matthewsalomon.wordpress.com/2008/04/03/dan-pagis-conversation/>

[2] For some aspects of the history of the digital humanities in Israel and their lack of what can be called a computational hermeneutics, especially in the context of Hebrew literature studies [Marienberg-Milikowsky 2019] [Marienberg-Milikowsky 2022].

[3] For a broader historical and theoretical background of the undogmatic principle of CATMA, see Flüh et al. 2021.

[4] We suggest the term *in-line annotation* to refer to the practice of annotating specific aspects in a text, rather than producing a label for the text as a whole.

[5] ViS-À-ViS will be able to interface with other annotation services in the future (such as INCEpTION, accessed 27 May 2023), at the moment it works exclusively with CATMA.

[6] <https://stereoscope.threedh.net>.

[7] ViS-À-ViS is a web-based tool using AmCharts library for Java Script; its front end is written in Vue.js while the back end is based on Python Flask. The data is imported through CATMA's Gitlab using a unique user access token. The data is then built in XML files using a hierarchy of text, tagset, tags and annotations. The development team of the tool consists of Dr. Oren Mishali, the tool's architect, Moshe Schorr and Yuval Erez, the programmers.

[8] For example, tools like Recogito (<https://recogito.pelagios.org>, accessed 27 May 2023) or INCEpTION (<https://inception-project.github.io>, accessed 27 May 2023). In addition, Münz-Manor and colleagues from the Technion, Israel Institute of Technology are currently training deep machine learning models, based on the manual annotations from the doctoral dissertation, to automatically detect metaphors in the medieval corpus of *piyyut* ("FigureOut - Automatic Detection of Metaphors in Hebrew Across the Eras", *DH2023*, Graz, Austria, poster session).

[9] This idea, we believe, has a lot of potential; we hope to discuss it in detail elsewhere.

Works Cited

Barthes 1975 Barthes, R. (1975) *The Pleasure of the Text* Translated by Richard Miller. New-York: Hill and Wang.

Culler 1992 Culler, J. (1992) "In Defence of Overinterpretation", in Colini S. (ed.) Umberto Eco, *Interpretation and Overinterpretation – with Richard Rorty, Jonathan Culler and Christine Brooke-Rose*. Cambridge: Cambridge University

Press.

- Drucker 2011** Drucker, J. (2011) "Humanities Approaches to Graphical Display", in *Digital Humanities Quarterly* 5(1).
- English 2010** English, J. F. (2010) "Everywhere and Nowhere: The Sociology of Literature After the 'Sociology of Literature'", in *New Literary History* 41(2), pp. v-xxiii.
- Flanders and Jannidis 2016** Flanders, J. and Jannidis, F. (2016) "Data Modeling", in Schreibman, S., Siemens, R., Unsworth, J. (eds.) *A New Companion to Digital Humanities*, Malden: Wiley Blackwell, pp. 229-237.
- Fleming 2017** Fleming, P. (2017) "Tragedy, for Example: Distant Reading and Exemplary Reading (Moretti)", *New Literary History* 48, pp. 437-455.
- Flüh et al. 2021** Flüh, M., Horstmann, J., Jacke, J., Schumacher, M. (2021) "Introduction: Undogmatic Reading - from Narratology to Digital Humanities and Back", in Flüh, M., Horstmann, J., Jacke, J., Schumacher, M. (eds.) *Toward Undogmatic Reading: Narratology, Digital Humanities and Beyond*, Hamburg: Hamburg University Press, pp. 11-29.
- Gius et al. 2022** Gius, E., Meister, J. C., Meister, M., Petris, M., Bruck, C., Jacke, J., Schumacher, M., Gerstorfer, D., Flüh, M., Horstmann, J. (2022) *CATMA 6 (Version 6.5)* Zenodo. DOI: 10.5281/zenodo.1470118.
- Goldstone and Underwood 2014** Goldstone, A., Underwood, T. (2014) "The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us", in *New Literary History* 45:3, pp. 359-384.
- Hammond 2017** Hammond, A. (2017) "The Double Bind of Validation: Distant Reading and the Digital Humanities' 'Trough of Disillusionment'", in *Literature Compass* 14:e12402.
- Herrnstein-Smith 2016** Herrnstein-Smith, B. (2016) "What Was 'Close Reading'?": A Century of Method in Literary Studies", in *Minnesota Review* 87, pp. 57-75.
- Horstmann 2020** Horstmann, J. (2020) "Undogmatic Literary Annotation with CATMA". *Annotations in Scholarly Editions and Research*, pp.157-176.
- Jacke 2018** Jacke, J. (2018) "Manuelle Annotation", in *forTEXT. Literatur digital erforschen*. <http://fortext.net/routinen/methoden/manuelle-annotation>.
- Jockers and Underwood 2016** Jockers, M. L., Underwood, T. (2016) "Text-Mining the Humanities", in Schreibman, S., Siemens, R., Unsworth, J., (eds.) *A new Companion to Digital Humanities*, Malden: Wiley Blackwell, pp. 291-306.
- Jänicke et al. 2015** Jänicke, S., Franzini, G., Scheuermann, G., Cheema, M. S. (2015) "On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges", in Borgo, R., Ganovelli F., and Viola I. (eds.) *Eurographics Conference on Visualization*.
- Kleymann and Strange 2021** Stereoscope. (2021) Hermeneutic Visualization in Literary Studies. Available at <https://stereoscope.threedh.net>.
- Marienberg-Milikowsky 2019** Marienberg-Milikowsky, I. (2019) "Beyond Digitization: Digital Humanities and the Case of Hebrew Literature", in *Digital Scholarship in the Humanities* 34:4, pp. 908-913.
- Marienberg-Milikowsky 2022** Marienberg-Milikowsky, I. (2022) "It Functions, and that's (almost) All: Tagging the Talmud", in Fiorimonte, D., Ricaurte, P., Chahudhuri, S. (eds.) *Global Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press, pp. 141-150.
- Marienberg-Milikowsky et al. 2022** Marienberg-Milikowsky, I., Vilenchik, D., Krohn, N., Kenzi, K., Portnikh, R. (2022) "An Experimental Undogmatic Modelling of (Hebrew) Literature: Philology, Literary Theory, and Computational Methods", in Horstmann J., Fischer, F. (eds.) *Digital Methods in Literary Studies - Special Issue of Textpraxis. Digital Journal for Philology* 6.
- Meister 2014** Meister, J. C. (2014) "Toward a Computational Narratology," in Maristella, A., Tomasi, F. (eds.) *Collaborative Research: Practices and Shared Infrastructures for Humanities Computing*, Padova: CLEUP, pp. 17-36.
- Moretti 2020** Moretti, F. (2020) "Simulating Dramatic Networks." *Journal of World Literature* 6, 1, 24-44, Available From: Brill <https://doi.org/10.1163/24056480-20201001> [Accessed 28 June 2023]
- Münz-Manor 2006a** Münz-Manor, O. (2006) "*Studies in Figurative Language of Pre-Classical Piyut*," PhD diss., Jerusalem: Hebrew University of Jerusalem.
- Münz-Manor 2006b** Münz-Manor, O. (2006) "Structural Ornamentations and Figurative Language in the Ancient Piyut," in *Jerusalem Studies in Hebrew Literature* 21, pp.19-38 (Hebrew)

- Münz-Manor 2009** Münz-Manor, O. (2009) “As the Apple Among Fruits, so the Priest When He Emerges’: Poetic Similies in Pre-Classical Poems of the ‘How Lovely’ Genre,” in *Ginzei Qedem - Genizah Research Annual* 5, pp. 165-188 (Hebrew)
- Münz-Manor 2011** Münz-Manor, O. (2011) “Figurative Language in Early Piyyut,” in Katsumata, N., van Bekkum, W. (eds.) *Giving a Diamond: Essays in Honor of Joseph Yahalom on the Occasion of His Seventieth Birthday*, Leiden: Brill, pp. 51–68.
- Münz-Manor 2016** Münz-Manor, O. (2016) “The Payytanic Epithet and its Connections with Figurative Language,” in *Jerusalem Studies in Hebrew Literature* 28, pp. 93-112 (Hebrew)
- Münz-Manor 2022** Münz-Manor, O. (2022) “Quantifying Piyyut: Computerized Explorations of Ancient Hebrew Poetry,” in *Jewish Studies Quarterly* 29, pp. 370-388.
- Münz-Manor et al 2020** Münz-Manor, O., Schorr, M., Mishali O., Kimmelfeld, B. (2020) “ViS-À-ViS : Detecting Similar Patterns in Annotated Literary Text,” in *VIS4DH*, <https://arxiv.org/abs/2009.02063>.
- Piez 2010** Piez, W. (2010) “Towards Hermeneutic Markup: An architectural Outline”, in *Digital Humanities 2010 Conference Abstracts Book*.
- Piper 2020** Piper, A. (2020) *Can We be Wrong? The Problem of Textual Evidence in a Time of Data* (Cambridge Elements: Digital Literary Studies), Cambridge: Cambridge University Press.
- Plank 2022** Plank, B. (2022) “The ‘Problem’ of Human Label Variation: On Ground Truth in Data, Modeling and Evaluation”, <https://arxiv.org/abs/2211.02570>.
- Rieder and Röhle 2012** Rieder, B. and Röhle, T. (2012) Digital Methods: Five Challenges. En D. M. Berry (Ed.), *Understanding Digital Humanities*. Palgrave Macmillan UK, pp. 67-84. https://doi.org/10.1057/9780230371934_4.
- Ross 2014** Ross, S. (2014) “In Praise of Overstating the Case: A review of Franco Moretti, Distant Reading”, *Digital Humanities Quarterly* 8(1).
- Sinclair and Rockwell 2016** Sinclair, S., Rockwell, G. (2016) Voyant Tools. Web. <http://voyant-tools.org/>.
- Swartz and Yahalom 2005** Swartz, M., Yahalom, J. (2005) *Avodah: An Anthology of Ancient Poetry for Yom Kippur*, University Park: Pennsylvania State University Press.
- Thomasson 2022** Thomasson, Amie, “Categories”, *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), Edward N. Zalta and Uri Nodelman (eds.), <https://plato.stanford.edu/archives/win2022/entries/categories/>.
- Underwood 2017** Underwood, T. (2017) “A Genealogy of Distant Reading”. *Digital Humanities Quarterly*, 11(2).
- Weingart 2022** Weingart, S. B. (2022) “Networked Society: The Moral Role of Computational Research in a Data-Driven World,” <http://scottbot.net/2014/09/> (accessed 22 June, 2022)



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.