Corpus-Driven Visualization of Textual Worlds

Jens Bruun Kofoed
Fjellhaug International University College,
Leifsgade 33.6, 2300 København S
jkofoed@fjellhaug.no

Abstract: EuroPLOT resources [http://resources.3bmoodle.dk] is a unique collection of close to 5000 images of people, places and customs related to the world of the Bible. Pictures selected for the database are high resolution images of 10 megapixels or more (except for images of historical interest). All pictures are tagged with information on the picture’s motif, relevant historical period(s), the name of the photographer, the date when the picture was taken, and Biblical references. The core vision is to use historical-archaeological pictures from primarily Israel-Palestine, but also the broader Biblical scene (Turkey, Lebanon, Syria, Jordan and Egypt) for visualization of textual worlds. The database is constructed to illustrate how a corpus-driven persuasive technology is able to make learning more effective and efficient through scaffolding. The database is integrated with the teaching tool Bible Online Learner, which is a tutor for the study of Hebrew and Greek based on the original Biblical texts, and relevant resources in the database are displayed as links in the Biblical text. The Flood account in Genesis 6-9 will be used as a case to demonstrate the potential of the database. The programmer of Bible Online Learner, Claus Tøndering will develop the automated search learning objects. The resource database will be enhanced with a mechanism that allows a resource (that is, a picture, a sound, a video, or a piece of text) to be associated with an entry in a Hebrew or Greek dictionary. Bible Online Learner will then be able to provide a link to that resource when the associated word occurs in the biblical text. We will furthermore for the project explore to what extent we can use internationally established standards for reference to location on the internet, i.e. a “Uniform Resource Name” and explore to what extent this will enable Bible OL to associate our dictionary entries with other digital library formats (e.g., Bible Odyssey, Wikimedia Commons, etc.) for canonical citation. The high resolution resources are made freely available for non-commercial uses, and therefore an important resource for Majority World researchers, teachers and learners. All resources will be tested in the classrooms in Copenhagen, Oslo, Mongolia, Madagascar, Kenya and elsewhere over the next years.

1. Introduction

Andreas Andreopoulos, in an introduction to icons, asserts that “a Greek who enters a Byzantine church in order to pray, or in order to participate in the liturgy, is very aware that he or she steps into an environment where everything, from the architecture of the building to the regularity of the censing, carries a specific theological ‘direction’ that is not distracting or simply decorative but, on the contrary, makes it possible for him or her to worship God fully and with a correct attitude” (Andreopoulos and O’Kane 2008, 84–85). By way of analogy, this is precisely what Bible Online Learner aims to achieve by adding images to the tailored support or “scaffold” offered to the learner of Biblical Hebrew and exegesis: By entering the Bible Online Learner “cathedral,” the student will soon be aware that he or she steps into an environment where everything, from the morphological and syntactical tagging of the underlying ETCBC4 database of the Hebrew Bible and the lexical and grammatical annotation of the Hebrew text to the iconographic ‘censing’ linked to its lexemes and verses,
carries pedagogical direction that makes it possible to fully – or at least adequately – understand the meaning of the Hebrew text!

2. Incentives for visualization of the Hebrew Bible

There are numerous reasons for tagging the Hebrew text in Bible Online Learner with links to images.

2.1. Analphabetism

The first and rather obvious reason is illiteracy. In the ancient Near East, where a large percentage of the population were functional analphabets, visual (and aural) learning was indispensable. And both Pre-Modern and Modern religious art reminds us, that images continued to be important for teaching less literate learners. Even the otherwise iconoclastic reformers made extensive use of images. Christoph Strohm, in an article on how the extensive use of Flugblätter – flyers, was made possible by the then recent invention of the printing press, describes how illustrations played a significant role in disseminating the new Gospel: “Illustrierte Flugblätter, die in dieser Weise wesentliche Anliegen auf den Begriff brachten und klare Alternativen aufzeigten, erfüllten in einer Welt, in der die Alphabetisierungsquote – zu mindest ausserhalb der Städte – niedrig war, eine wichtige Funktion” (Strohm 2015, 15). And though there are significant variations in literacy rates between world regions in the modern world, the 2013 UNESCO Institute for Statistics (UIS) data indicates adult literacy rates of only 78.12% in the Arab States, 67.55% in South and West Asia, and 59.76% in Sub-Saharan Africa (UNESCO Institute for Statistics 2015). And due to demographic changes caused by immigration and waves of refugees, literacy rates have actually decreased (albeit from a near 100% level) in North America and Europe in the past two decades. Ergo: There are still a lot of people who rely on pictures or material images because they can’t read.

The most important incentives for combining text and image in teaching and learning, however, come from cognition research.

2.2. Cognitive Theory

Neurological research has demonstrated that cognitive maps are generated by input from both the visual and aural system, in order to reduce cognitive load, enhance recall and learning of information (Kitchin 1994). Cognitive linguistics has shown, in the words of Martin Klingbeil, “that both iconography (the material image) and metaphor (the literary/verbal image) converge in the cognitive image domain,” and that “one can establish the semantics of material and literary images on a micro-level and then map these elements to one another” (Klingbeil 2014, 135–36; cf. Klingbeil 2009). In other words, verbal and material images are mutually explanatory. An observation that is corroborated by the so-called Cognitive Load Theory which holds that the use of visual symbols with verbalised explanations is more effective than either visual presentation (graphics + text) or verbal explanations alone (Tuovinen 2000).

2.3. Learning Style

Another and related incentive comes from pedagogical research on learning styles. When, 15 years ago, I was helping my then eight year old daughter with her math homework, she simply couldn’t understand my explanations and eventually cried out in frustration: “Dad, I cannot picture it in my mind!” At first I was puzzled by how math calculations could be “pictured” in the learning process,
but eventually I realized that what she was saying in her own language was: “Dad, I have a predominately visual learning style, but the workbook doesn’t support that.” And the reason why she didn’t understand my explanation was that I was actually copying the workbook’s non-visual teaching style.

Now, learning style is a heavily debated concept, and refer to a steadily growing number of competing and contested theories, but even if the critics are right in their skepticism and more randomized controlled trials are needed, there is widespread scholarly agreement that the problem is not that different individual learning preferences don’t exist, but that researchers struggle to identify and describe them in evidence-based theories. With these caveats in mind, both cognitive theory and data from tests of large cohorts of learners still suggest that, e.g., the learning style categories of David A. Kolb (accommodator = concrete experience + active experiment, converger = abstract conceptualization + active experiment, diverger = concrete experience + reflective observation, assimilator = abstract conceptualization + reflective observation), Neil Fleming’s VARK-model (Visual, Auditory, Read/Write, and Kinesthetic learning), or the grouping of more than thirty categories into two independent dimensions (wholist-analytic and verbal-imagery) by Richard J. Riding, I. Cheema and Stephen Rayner (Kolb 2015; Fleming 2001; Fleming 2016; Riding and Cheema 1991; Riding and Rayner 1998) reflect the reality of learning, and, more important all point to the importance of combining text and image in teaching and learning.

2.4. Images as an Inroad to Culture

A final incentive that must be mentioned is the fact that images are important inroads to a culture’s cognitive environment. Izaak J. de Hulster, e.g., asserts that

Cultural theorists and art historians emphasize that images, no less than texts, are a constitutive component of any given culture’s symbol system. Rather than serving as mere decorations, images - whether ancient, medieval, or modern - are capable of conveying crucial information between senders and receivers. ... Still further, as communicative media, images no less than texts can be described as cultural repositories, containing information about society, religions, politics, economics and so forth. As such, they reflect the thoughts of human agents (e.g., the artists and/or people who commissioned or purchased the images) and were designed to be understood - even ‘read’ - by their respective audiences. In this way, and exactly as textual remains, ancient Near Eastern iconography provides a window into the cultural, social, religious, and political world that lies behind the Hebrew Bible (Hulster 2015, 21).

3. Challenges to Visualisation

3.1. Copyright infringement

Whereas a decision has been made to provide open access to the more than 5000 high resolution and Biblical relevant images in the EuroPLOT resources database, copyright issues are much more complicated, of course, as far as links to external, copyrighted resources are concerned. Fortunately, however, a very fresh judgment from The Court of Justice of the European Union rules that “[t]he posting of a hyperlink on a website to works protected by copyright and published without the author’s consent on another website does not constitute a ‘communication to the public’ when the person who posts that link does not seek financial gain and acts without knowledge that those works have been
published illegally” (Court of Justice of the European Union 2016). In other words: As long as copyrighted images are not imported into the EuroPLOT resources database, but only made accessible through links tagged to the Hebrew text in Bible Online Learner, there is no infringement of copyright vis-à-vis copyright holders based in and/or covered by EU copyright legislation. This makes it possible, e.g., to link to the growing number of open access image databases published on the internet by leading European Museums such as The British Museum in London, Vorderasiatisches Museum in Berlin, The Louvre in Paris, etc. Whether this is true for copyright holders outside the European Union needs to be checked, of course, before links are provided to online, open-access, but copyrighted material from, e.g., SBL’s Atlanta-based Bible Odyssey, the New York-based Metropolitan Museum of Art, the San Francisco-based Pinterest, etc.

3.2. Unstable URLs

1. Another related challenge is the so-called Internet “link rot” or “404 page not found” errors due to renamed, moved or deleted internet content. And though considerable efforts have been made to secure URL stability through persistent identifiers (PIDs) such as Uniform Resource Locators (PURLs), Uniform Resource Names (URNs) or Digital Object Identifiers (DOI numbers) for academic papers, there is still a high risk that the digital rug will be pulled out from under them from time to time. The problem needs a global solution, of course, and until such a solution is in place, tagging with links needs to be made by observing a number of rules that minimize link rot, such as Kille’s advice to put in only essential links and to avoid linking to anything that might go away, e.g., personal or short-term project websites (Kille 2015).

3.3. Majority World

Having taught in Majority World countries like Madagascar and Mongolia, I am too well aware of how poor or non-existent access to the Internet impedes the possibilities of doing research

A representative, sub-Saharan case is Madagascar. Between 1990 and 2008, Madagascar was completely isolated from the Internet because of total lack of competition and a connection of very poor quality. Positive developments in the internet and broadband sector have begun following the arrival of the first international submarine fiber optic cables, LION (the Lower Indian Ocean Network) and EASSy (the East African Submarine Cable System) on the island in 2009 and 2010. This ended the country’s dependency on satellites for international connections, bringing down the cost of international bandwidth and making internet access more affordable to a wider part of the population. However, the reality on the ground is still grim. The prices have dropped, but in 2013 they were still exorbitant for the majority of the Malagasy. The average wages are approximately 100 to 200 dollars per month (in best of the cases) and an ADSL connection costs around 75 dollars per month. Being one of the sub-Saharan countries with the lowest score on the MPI, even state universities struggle with the costs of high-speed internet access. Add to this that when it comes to electricity, Africa remains the dark continent. There are a billion Africans, but they use only 4% of the world’s electricity. Most of that is round the edges, in Egypt, the Maghreb and South Africa. The rest of Africa is unlit because either because electricity is unavailable or too expensive.

Several initiatives have been taken by minority world agencies, institutions, and organizations to increase the availability and thereby also the dissemination of scholarly research in Africa. The LIBLICENSE homepage (updated on May 5, 2014), e.g., lists 44 current developing nations initiatives
One of the more significant examples is the 2008 JSTOR’s African Access Initiative, in which JSTOR waives the standard participation fee for any non-profit African institution (JSTOR 2016).

In addition to such efforts to make academic research available to majority world institutions, the production and sharing of learning objects is an important and less costly way to share the latest research, to expose majority world teachers/learners to best pedagogical practice, and to expose minority world teachers to majority world perspectives. Post-colonialism at its best, as it were!

As Bible Online Learner is an online resource itself, it will be of little help to Majority World students with poor or non-existent internet connections, if it were not for the fact that, e.g., Africa have seen strong growth in mobile broadband use and data traffic in recent years. The recently updated “Madagascar - Telecoms, Mobile, Broadband, and Digital Media – Statistics and Analysis” report from BuddeComm estimates the penetration of fixed Internet to 3.5% and Mobile SIM to 40% (BuddeComm 2016). As the penetration estimate for Mobile SIM is 40% of the population, the percentage of Malagasy students with Mobile SIM is no doubt considerably higher, and as first generation cell phones are gradually replaced by smartphones, most Malagasy students will soon have mobile access to the Internet. In order to overcome the problem of poor or lacking fixed internet access, the developers of Bible Online Learner have therefore been dedicated to design the platform as responsive web in order for Majority World smartphone and tablets with SIM-card users to access the learning material.

3.4. Iconographic Exegesis

A final – and very different – is the need for methodological awareness in the use of images to visualize the Hebrew text. Reading texts and seeing images are two very different cognitive processes, and both biblical and iconographic exegesis have their own distinct exegetical steps. Fortunately, a steadily growing number of publications on ancient Near Eastern iconography has appeared in the wake of the pioneering work of Othmar Keel and the so-called Fribourg School, providing us with a solid point of departure for analysis of images from or relating to the world(s) that produced the biblical text. The inductive approach taken by Othmar Keel, Christoph Uehlinger and other proponents of the Fribourg School, has now been supplemented by more deductive ones. Izaak J. de Hulster, e.g., defines iconographic exegesis as “a historical method … intended to illuminate biblical texts (their contents, concepts, ideas, etc.) and their historical contexts (religious, social, cultural, political, economic) with the help of images in order to understand the text as much as possible in the way the first receiving communities did” (Hulster 2009, 20). The first challenge is, of course, how to find relevant images for illumination of the selected Biblical text or theme, and Hulster points to three general aspects of the image-text relationship that may serve as a useful guide in the selection of images:

1. **Image-text congruence**: Which images and texts can be thought of as being related, and to what extent do they share similar (or manifest different) themes, motifs, and/or subject matter?

---

1 According to http://liblicense.crl.edu “the LIBLICENSE project has been hosted since November 2011 by the Center for Research Libraries. Founded in 1949, the Center for Research Libraries is a consortium of over 250 academic and independent research libraries in the U.S., Canada and Hong Kong. CRL supports advanced research and teaching in the humanities, sciences, and social sciences by preserving and making available to scholars primary source materials critical to those disciplines.”
2. **Image-text correlation**: At what level are images and texts related, and how have scholars understood both the type and direction of interaction that occur between these two media?

3. **Image-text contiguity**: To what extent does the presence of historical lines of influence and/or mechanisms of contact determine whether at given image and text are thought to be related, and what are the implications for comparative methodologies? (Hulster 2015, 23)

Once the images relevant for illuminating a given biblical text or theme has been identified, Hulster, on the basis of, not least, E. Panovsky (Panovsky 1932; Panovsky 1955), but also Oskar Bätschmann (Bätschmann 2001), and R. van Straten (Straten 1994), presents the following summary of Panofsky’s approach to image analysis and argues that it “remains the most important starting point for methodological reflections’ in iconographic exegesis”(Hulster 2009, 70; Hulster 2015, 38):

<table>
<thead>
<tr>
<th>Interpretive Object</th>
<th>Interpretive Task</th>
<th>Interpretive method</th>
<th>Interpretive Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary or natural subject matter: the world of artistic motifs</td>
<td>Pre-iconographical description of motifs</td>
<td>Practical experience: familiarity with objects/motifs and events</td>
<td>History of style: insight into why objects and events were expressed by specific forms at certain times</td>
</tr>
<tr>
<td>Secondary or conventional subject matter: the world of images and image complexes (stories, allegories)</td>
<td>Iconographical analysis in the narrow sense</td>
<td>Knowledge of literary sources: familiarity with themes and concepts at work in an image</td>
<td>History of types: insight into why specific themes or concepts were expressed by objects and events at certain times</td>
</tr>
<tr>
<td>Intrinsic meaning or content: the world of “symbolical” values reflected/present in the work</td>
<td>Iconographical analysis in the deeper sense or, better, what might be called iconographical interpretation/synthesis (“iconology”)</td>
<td>“Synthetic intuition”: familiarity with tendencies of the human mind</td>
<td>History of cultural “symbols”: insight into why tendencies of the human mind were expressed by specific themes and concepts at certain times</td>
</tr>
</tbody>
</table>

Hulster’s declared goal is to provide a methodology for iconographic exegesis of images relevant for the understanding of biblical texts, and Hulster has therefore been criticized for maintaining a superiority of the biblical text and restricting image analysis to an auxiliary discipline of biblical exegesis (Schurte 2010). This is not a necessary critique, however. As long as images (in the same way as archaeological data) are treated as sources on the same level as the biblical text and being analysed with source-specific analytical methods, images may be used to illuminate all sorts of things, including biblical texts. Critics are right, of course, that deductive methods (like the one recommended by Hulster) must be supplemented by inductive ones (as implemented by the Fribourg School), but there is no methodological flaw in using both approaches as long as images are “exegeted” on their own terms before they are brought to bear on texts. On the contrary, dialectical use of both inductive and deductive methods may even lead to better understandings of images and thereby provide the exegete with a more substantial sidelight in his efforts to understand the biblical text.
4. Bible Online Learner

Bible Online Learner was – and continues to be - developed primarily as a persuasive technology tool to enhance language learning based on the corpus of the Hebrew Bible – the co-called ETCBC4 Database – created and maintained by the Eep Talstra Center for Bible and Computer in Amsterdam. It is “corpus-driven” in order to stimulate curiosity from day one, not only in learning Hebrew, but also in the *interpretation* of the Hebrew text. The feature of tagging both lexemes and verses with links to pictorial (and/or textual) information is a means to the same end.

The text is tagged with links to three types of images:

1. Contemporary images relevant for exegesis – but not intentional interpretations of – the biblical text.
2. Contemporary or later conscious pictorial interpretations of the biblical text.
3. Modern photos of sites, items or customs mentioned in the biblical text.

The three categories serve different and only partially overlapping exegetical purposes. The first category of images has the advantage of providing insight into the cognitive environment of the culture that produced the text. The second category throws light, primarily, on the reception history of the text, whereas the third category, e.g., may enhance the understanding of geographical, topographical and ethnographic features in the text.

Links, as illustrated in the figure below, can be either internal or external to the (green) Bible Online Learner resources family.

It is also possible to create lexeme-tags with links to other types of relevant media such as music, film, Logos Bible Software resources etc. As for Logos, it is possible to link to a search result in Logos. The link `https://ref.ly/logos4/Search;kind=Media;q=Ararat;match=stem`, e.g., starts the Logos
Bible Software application and reproduces the search result for “Ararat” in “all media in all resources.” The result depends, of course, on whether Logos Bible Software is installed, and, if that is the case, which resources are available in the particular Logos installation. And though I haven’t tried it, I take it that the same is possible in relation to other Bible Software products like Accordance, BibleWorks, and others.

4.1. The Flood Story in Genesis 6-9 as Example

4.1.1. Lexeme tags

As first category images relevant for enhancing the understanding of the Flood story I have chosen to tag the lexeme מַבּוּל ‘flood’ (e.g., Gen 7:6) with a link to a picture of the Gilgamesh Tablet from the British Museum Online Collection.² The link is not to the picture itself, but to the page containing the picture, since the curator’s comments are necessary in order for the uninformed to understand the relevance of the picture for the interpretation of the biblical Flood Story.

A second example is the lexeme תֵּבָה ‘ark’ (e.g., Gen 8:4), which is tagged with two links. The first leads to a blog entry with a 1:5 scale reproduction of a coracle ark built to the specifications within the so-called Ark Tablet describing the Mesopotamian Flood hero Utnapishtim’s escape from the Flood.³ The second link is pointing to a lengthy article on the Mesopotamian Flood stories including a picture of the so-called Ark Tablet.⁴

The third lexeme, תְּהוֹם ‘deep’ (e.g., Gen 8:2) is tagged with a link to the entry on the freshwater god Enki/Ea (god) in the Ancient Mesopotamian Gods and Goddesses (AMGG) list of deities in the University of Pennsylvania’s Open Richly Annotated Cuneiform Corpus (ORACC) project.⁵ A second link refer the user to a picture from The British Library’s collection showing a bearded god, probably Ninurta, fighting the dragon-monster Tiamat.⁶ None of these images are conscious interpretations of the Biblical text but they all inform us about the conceptual world that produced the Biblical text. And as far as the last example is concerned: Even though the Biblical lexeme tehom is not etymologically derived from the Akkadian proper noun Tiamat – they are in all likelihood both derived from the same common etymological ancestor – the image of a god fighting a freshwater dragon-monster informs us about how some – probably the majority of – ancient Near Easterners thought about the underground water.

---

⁵ http://oracc.museum.upenn.edu/amgg/listofdeities/enki.
⁶ http://www.britishmuseum.org/research/collection_online/collection_object_details/collection_image_gallery.aspx?partId=1&assetId=417164001&objectId=277961
As an example on contemporary or later conscious pictorial interpretations of the biblical text, I have chosen to tag the lexeme נֹחַ ‘Noah’ (e.g., Gen 8:1) with links to fresco paintings from the catacombs in Rome. The link is to the picture itself, and since there is no information associated with the picture, the link needs either to be supplemented by a second link with a description of the picture or used in a “guided” class room discussion on, e.g., why the ark is modelled as a (wooden) chest, and why Noah is posed with his arms lifted.

A second example is a “moving pictures” link to a Youtube video on Los Angeles Opera performance of Benjamin Britten’s opera Noy’s Fludde (Noah’s Flood) in which the stage director explains why the Flood Story is still important to tell.8

As third category image, I have chosen to tag the lexeme אֲרָרַט ‘Ararat’ with a link to a map showing the approximate size of ancient Urartu (Ararat).

The maximum number of lexeme-tags is three, but if, e.g., one needs to link to a collection of pictures, this can be done by linking to an external resource-page containing the pictures themselves or with links to them. By linking to external resource-pages, Bible Online Learner may be used as an integrated part of password-protected or paywall-restricted courses on Learning Management Systems such as Blackboard, Moodle etc. The lexeme אָמַר ‘to utter, say’ is tagged, e.g., with a link to Nicolai Winther-Nielsen’s Bereshit Basic Biblical Hebrew Moodle course with guest access to the first 12 sessions.

4.1.2. Verse tags

Verse tags are created, not in Bible Online Learner, but in the EuroPLOT Resources administrative interface, where images are tagged with all relevant biblical references. And if there are pictures in the EuroPLOT Resources database tagged with a reference to a given verse, an icon at the beginning of the verse in Bible Online Learner will indicate that there are picture resources available, and by clicking the icon all pictures tagged with the particular reference will be shown. Figure x, e.g., shows the results for Gen 8:4, where the “mountains of Ararat” – ancient Urartu – are mentioned.

4.2. Dreams and Plans

4.2.1. Course Specific Link sets

7 https://vhoagland.files.wordpress.com/2015/06/moses-noah.jpg.
8 https://www.youtube.com/watch?v=RqKezabmMTM&t=60s
The links referred to above may or may not be very interesting or relevant for you, but the next step in our development plans for Bible Online Learner is to make it possible for users to create course specific link sets. And it would be very much in the spirit of the developer team, if teachers decided to *share* such links sets with the rest of the world, not least the Majority World, where access to updated teaching and learning material is limited.

4.2.2. Tags to Phrases and Clauses

Another possibility we will be looking into is the possibility to tag not only on lexeme- and verse-level, but also on phrase- and clause-level.

4.2.3. Automatically Generated Links to External Databases

In our dreams, we also imagine automatically generated links to external databases. As already mentioned, Bible Online Learner automatically searches the EuroPLOT image database for pictures, and this could just as well be done with other databases. The Bibel+Orient Datenbank Online, e.g., makes extensive use of bible references in its *Darstellungen*, that is, the texts describing the images, and it would be extremely useful to have an “on/off switch” for links to such databases as well as the one already present for the EuroPLOT database. Still dream, yes, but sometimes dreams come true!

5. Bibliography


Tuovinen, Juhani E. 2000. ‘Cognition Research and Educational Multimedia.’ Seminar paper, Centre for Learning and Teaching Support, Monash University, Australia.