Bible Online Learner:
Technical Documentation

Claus Tøndering

20 March 2015

Contents

Chapter 1

Introduction

TODO: Describe how to set up a running environment: git download, MySQL configuration, MQL support, Google login key.

You must read this chapter.

This document gives a detailed technical description of the internal workings of Bibel Online Learner (Bible OL). The document is intended for people who need to install Bible OL on their own server and developers who need to modify or expand the way Bible OL works.

1.1 License and Copyright

1.2 Terminology

Unfortunately, various parts of the system do not use a completely uniform terminology. This section gives an overview of what you may come across here:

Term	Meaning
quiz exercise	These terms are synonymous. They refer to an exercise executed by a user.
quiz template	A file located in the `quizzes` directory. In XML form it describes how Bible OL should generate questions for an exercise.
question question item	An exercise consists of several questions (see Figure 1.1). Each question contains several question items.
sentence unit quiz object question object	These terms are synonymous. The terms refer to the Emdros objects that are the subject of a question item. They are marked in red in the question text. (See Figure 1.1.)

(Continued...)

[1] http://opensource.org/licenses/MIT

[2] http://creativecommons.org/licenses/by/4.0/

Term	Meaning
display feature	An Emdros feature presented to the user as part of a question item (see Figure 1.1).
request feature	An Emdros feature which the user must provide as part of a question item (see Figure 1.1).
passages universe	The collection of Bible passages from which an exercise draws its sentences. The term "universe" is found in older parts of the code.
grammar selection box	A box found on the left part of the display where the user can select what in-line grammatical features to display. See Figure 1.2.
grammar information box	A box found on the right part of the display containing grammatical information about the word under the mouse pointer. See Figure 1.2.
grammar hierarchy	The organization of components of text. At the lowest level of the grammar hierarchy we have the *words*. Above that we may find *phrases*, then *clauses*, and finally *sentences*. The exact words for the levels of the grammar hierarchy varies between databases.

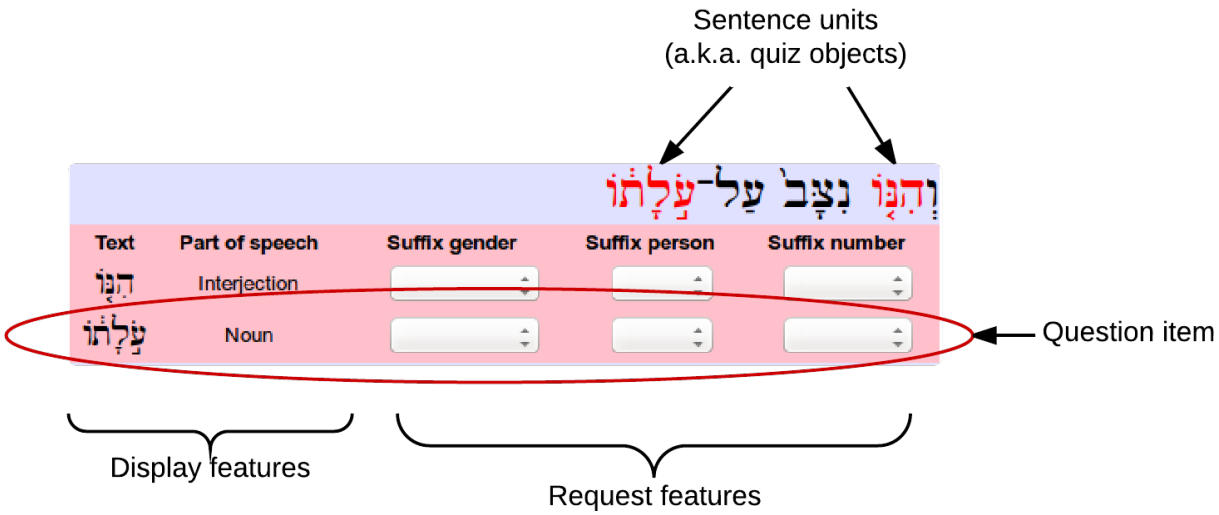


Figure 1.1: A question containing two question items.

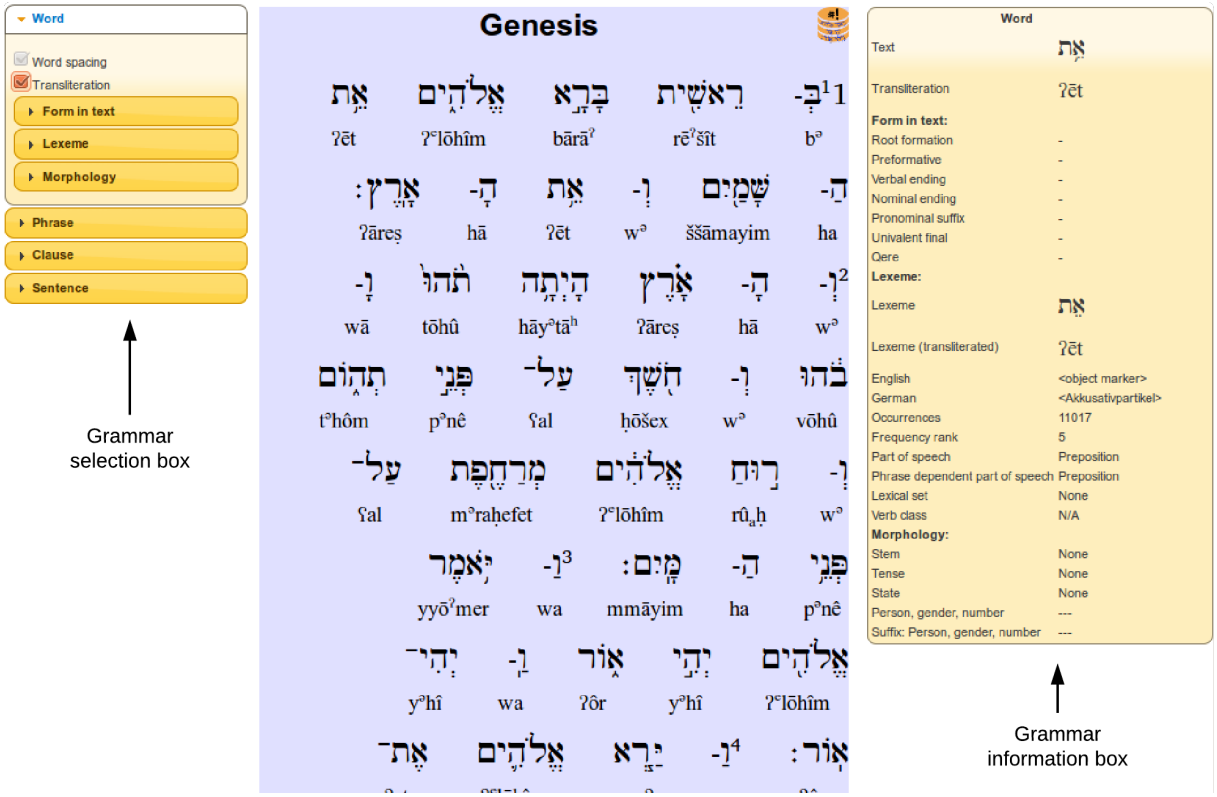


Figure 1.2: The grammar selection box and the grammar information box.

Chapter 2

History

Read this chapter if you want to.

2.1 EQG: A Java Applet (2008)

In 2008 professor Nicolai Winther-Nielsen told me about a text database system, Emdros, developed by Ulrik Sandborg-Petersen. Nicolai was teaching Biblical Hebrew at what was then the Copenhagen School of Theology.[1]

I further learned that Emdros databases exist with the entire text of the Bible in the original languages, complete with grammatical information about every single word.

We discussed how these tools could be used by Nicolai in his teaching, and in the autumn of 2008 the first version of *EQG*, the *Emdros-based Quiz Generator*, was demonstrated, running as a Java applet in a web browser.

2.2 3ET: A Stand-alone Java Program (2009-2010)

The Java applet solution was not practical, especially since there was no easy way to access an Emdros database over a network connection. In 2009 the applet was therefore abandoned in favour of a stand-alone PC application, still written in Java but running directly under Microsoft Windows.

The name was changed to *3ET, Ezer's Emdros-based Exercise Tool*, and an attempt was made to market it through 3BM, a company owned by Nicolai and his colleague Jens Bruun Kofoed.

The name 3ET is still reflected in the file extension `.3et` used in exercise files.

2.3 PLOTLearner: A EuroPLOT Product (2011-2013)

In 2011 3ET became part of an EU project about *Persuasive Learning Objects and Technologies*, PLOT. The project became known as EuroPLOT,[2] and 3ET changed its name to *PLOTLearner*.

Since PLOTLearner was part of an EU project the source code was made open under an MIT License.[3] PLOTLearner was the last Java-based version of the product, and it can be downloaded from `http://eplot.3bmoodle.dk/index.php/downloads`.

[1] A.k.a. Dansk Bibel-Institut. This later became the Fjellhaug International University College Denmark.
[2] `http://www.eplot.eu`.
[3] `http://opensource.org/licenses/MIT`.

2.4 Bible Online Learner: Web-based (since 2013)

In 2013 I started moving PLOTLearner from a Java-based stand-alone PC application to a web-based solution. Once again, the name of the product was change once more and became *Bible Online Learner*, or *Bible OL* for short. This is product that is described in this document.

Chapter 3

Programming Languages and Frameworks

You must read this chapter.

A considerable number of programming languages and other specification languages are used in the creation and execution of Bible OL. This chapter gives a brief overview of these languages and points you to where you may learn more about them.

3.1 PHP

The main language used on the server side is PHP[1]. PHP is a popular general-purpose scripting language that is especially suited to web development.

In order to execute Bible OL, the PHP implementation on the server must be enhanced with features to execute MQL commands. This is described in Section 6.1.1.

3.2 CodeIgniter

Bible OL uses a PHP framework known as *CodeIgniter.*[2]

More information is given in Section 15.

3.3 SQL

SQL[3] is a language for manipulating a relational database. Bible OL uses the MySQL database[4] to store information about users who have an account on the Bible OL website.

SQL commands are executed from PHP code through CodeIgniter.

3.4 MQL

MQL is a language for manipulating Emdros[5] text databases. The PHP implementation on the server must be extended with function to execute MQL commands.

MQL and Emdros are described in greater detail in Chapter 5.

[1] http://php.net.
[2] http://www.codeigniter.com.
[3] See https://en.wikipedia.org/wiki/SQL.
[4] http://www.mysql.com.
[5] http://emdros.org.

3.5 HTML

The generated web pages use HTML version 5.[6]

3.6 CSS

CSS (Cascading Style Sheets)[7] is a language for specifying the layout style of a web page. However, only a small part of the Bible OL styles are written directly in CSS. Most styling is written in Less which is then compiled into CSS.

3.7 Less

Less[8] is a CSS pre-processor, meaning that it extends the CSS language, adding features that allow variables, mixins, functions and many other techniques that allow you to make CSS that is more maintainable, themable and extendable.

Although Less style files can be compiled when used in a browser, the Bible OL implementation compiles Less files only once and stores the resulting CSS files.

More information about how Less is used is given in Chapter 18.

3.8 JavaScript

On the client side (that is, in the user's browser) the software is loaded as JavaScript[9] code. However, only a small part of Bible OL is written directly in JavaScript. Most client-side software is written in TypeScript which is then compiled into JavaScript.

3.9 TypeScript

TypeScript[10] is a superset of JavaScript that adds strong typing and proper classes to JavaScript.

Most of the client-side software of Bible OL is written in TypeScript which is then compiled into JavaScript.

More information about how TypeScript is used is given in Section 19.1.

3.10 JSON

JSON[11] is a text-based data-interchange format. It is used to transfer data between the server and the client.

3.11 jQuery and jQuery UI

On the client side Bible OL uses a JavaScript/TypeScript framework known as *jQuery*[12] and its associate user interface functions *jQuery UI*.[13]

[6] See https://en.wikipedia.org/wiki/HTML5.
[7] See https://en.wikipedia.org/wiki/Cascading_Style_Sheets
[8] http://lesscss.org.
[9] See https://developer.mozilla.org/en-US/docs/Web/JavaScript.
[10] http://www.typescriptlang.org.
[11] http://json.org.
[12] http://jquery.com.
[13] http://jqueryui.com.

3.12 What You Must Know

If you plan to modify the Bible OL server-side code, you must know how to program in PHP, and you must understand the CodeIgniter framework. You may also need to have a good understanding of HTML, Less, CSS, SQL, MQL, and JSON.

If you plan to modify the Bible OL client-side code, you must know how to program in TypeScript, and you must understand the jQuery framework and the jQuery UI functions. You may also need to have a good understanding of HTML, CSS, JavaScript, and JSON.

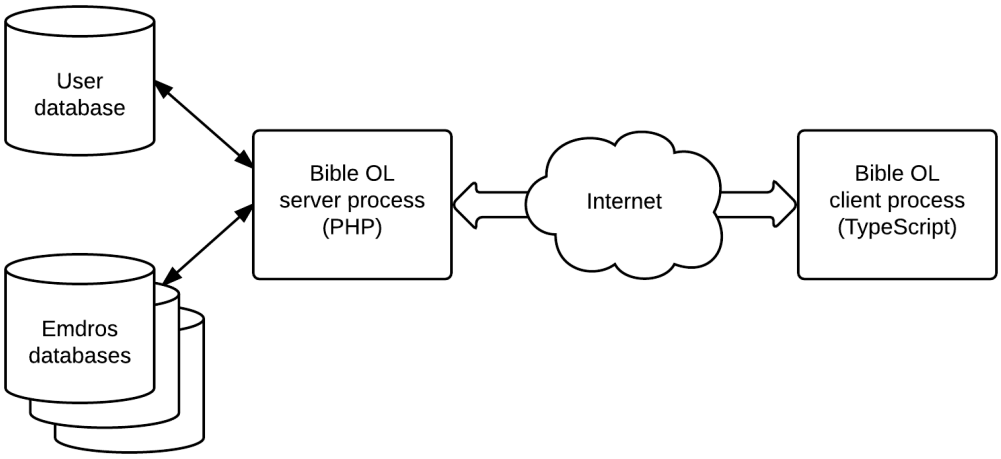
Obviously, you also need a good understanding of how Bible OL works from a user's perspective.

Chapter 4

High-level System Architecture

You must read this chapter.

Bible OL consists of two main components, the server and the client, as shown in the following illustration.



The server process runs on a Linux computer, but can probably quite easily be ported to a Windows server that supports PHP, MySQL, and Emdros. It is programmed in PHP and has access to a number of databases:

- The user database, which contains information about registered users.
- A number of Emdros databases, which contain the text and grammatical information for the Old and New Testaments (and potentially other texts as well).

The client is a web browser that accesses the server over an internet connection. It requests the server to provide an exercise or a portion of a text, which it then displays in the browser window. The client code is primarily written in TypeScript, which has been translated to JavaScript so that the browser can execute it.

Most of the layout is done in the client. The server generates HTML code for the main items in the window, but the actual text to be displayed is stored in variables in the client code, and the layout of this information is performed by the client.

The data exchange between the client and the server is either a two-step or a three-step exchange. If the client merely requests a text to be displayed, the data exchange is:

1. Client requests the server to load a particular URL. The text to display is coded into this URL.
2. The server sends the requested text to the client.

If the client requests an exercise to be displayed, the data exchange is:

1. Client requests the server to load a particular URL. The URL contains information about which quiz template to use and how many questions to generate.
2. The server sends the requested exercise to the client.
3. When the user click on the "Finish" button, the user's answers are sent to the server.

Note that there is no data exchange between the client and the server when the user presses the "Check answer" or "Show answer" buttons. These buttons execute code that is local on the server. (This means, that in theory a student can find the correct answers to the questions by looking at the source code for the web page. This is, however, quite difficult to do, and it is not considered a serious flaw in the design since it would require considerable knowledge on the part of the student.)

Chapter 5

Understanding Emdros

You must read the first section of this chapter. If you need to use MQL, you should read the whole chapter and also some of the documentation that comes with Emdros.

Emdros is a database system for storing and retrieving annotated text. Emdros was developed by Ulrik Sandborg-Petersen. A short introduction to Emdros for linguists is found at http://emdros.org/petersen-emdros-COLING-2004.pdf. More documents are available at http://emdros.org/docs.html.

Here, only a very brief description of the system will be given. Emdros divides a text into *objects*. Typical objects are *word, clause,* and *sentence*. For biblical texts, objects such as *book, chapter,* and *verse* are also used. Each occurrence of the smallest object, typically a word, is identified by a number, called a *monad* in Emdros terminology.

As an example, consider this text: "The boy, who had red hair, was sitting on the floor." This sentence consists of 11 words. We can also identify two clauses, "The boy...was sitting on the floor" and "who had red hair". Emdros assigns a monad (a positive integer) to each word object:

Monad	Word
1	The
2	boy,
3	who
4	had
5	red
6	hair,

Monad	Word
7	was
8	sitting
9	on
10	the
11	floor.

Additionally, sets of monads are used to identify clause objects: Monads { 1-2, 7-11 } identify one clause, and monads { 3-6 } identify another. Finally, the monads { 1-11 } identify a sentence object.

Associated with each object are a number of *features* that describe various characteristics of the object. For a word object, typical features could be *part of speech, gender, number, tense,* and *mood.* In the above sentence, word object 4 (the word "had") could, for example, have these features:

Feature	Value
Text	"had"
Part of speech	Verb
Tense	Past
Number	Singular
Person	3rd
Lexeme	"have"

Similarly, the two clause objects could have a feature called *type* with the values *main* and *subordinate,* respectively.

The exact set of objects and features available in a database is entirely up to the person who creates the database.

In addition to the monads, Emdros objects can also be identified by an ID_D. Like a monad, the ID_D is an integer, but every object in the Emdros database has a unique ID_D. So a single ID_D may refer to a single word or a clause or a sentence. In the above example, the word "who" with monad 3 may have ID_D=8, and the clause "who had red hair" with monads { 3-6 } may have ID_D=12.

Emdros comes with a query language, called MQL[1], that allows a user to search a corpus for objects with various features. MQL queries can be quite simple, such as "find all verbs in the past tense," or very complex, such as "find all sentences containing a singular pronoun, followed by at most three words, followed by a verb in the present tense, except cases where the verb is derived from 'to be'."

The exact syntax of MQL queries can be quite arcane and is beyond the scope of this paper, but examples can be found in http://emdros.org/petersen-emdros-COLING-2004.pdf.

[1] Mini Query Language. (Quite a misnomer since this is a very powerful language.)

Chapter 6

Development and Runtime System Setup

Read as much of this chapter as you find necessary.

This chapter describes how to set up a complete development system for working with all aspects of Bible OL development. Depending on the type of development you are going to do, you may not need all of this.

This chapter also describes how to set up a runtime system; this will be a subset of a development system.

Much of the description here is quite vague. In many cases I am simply describing what *I* have done. Your system may be different, and you may need to do things I have not described.

For Bible OL, I use a computer with the Linux operating system (the Ubuntu distribution). I am sure the system could also be set up on a Windows computer, but I have not tried it and you will not find any instructions how to do it here.

The following should be installed:

- Apache (web server). I use version 2.4.7, but version 2.2.22 also works.
- MySQL (database system). I use version 5.5.41.
- SQLite3 (database system used by Emdros). I use version 3.8.2.
- PHP. I use version 5.5, but version 5.3 also works.
- (Development system only:) nodejs (JavaScript runtime, used for compiling TypeScript and Less). I use version 0.10.25.
- (Development system only:) npm (package manager for nodejs). I use version 1.3.10. This package may not strictly be required, but it is useful for installing nodejs modules.
- (Development system only:) git. I use version 1.9.1.

All of this can be installed on an Ubuntu system using the usual *apt-get* installation program. I have not done a thorough investigation into the minimum required versions of these software packages; you should not see the versions mentioned above as minimum requirements.

You will also need to install the Emdros database system. On the development system you will also need a Less compiler and a TypeScript compiler. This is detailed in the following sections.

6.1 Installing Emdros

6.1.1 Emdros on a Development System

I use Emdros version 3.4.0.pre03, a pre-release; but versions 3.4.0 should probably be fine as well. Emdros can be downloaded from http://sourceforge.net/projects/emdros/files/emdros.

Follow the instructions that come with Emdros to compile and install it on your computer. You will need a C++ compiler and the Gnu make system (and probably a few other things as well – check the Emdros documentation).

When you compile Emdros, be sure to include support for PHP. I use this *configure* command to set up the Emdros compilation:

```
./configure --with-sqlite3=local --with-mysql=yes --with-default-backend=sqlite3
    --with-swig-language-php=yes --with-swig-language-java=no --with-swig-language-python=no
    --with-swig-language-perl=no
```

If you want to include support for Java, Python, or Perl, you will need to change this command.

When you have compiled and installed Emdros (using the *make* and *sudo make install* commands, respectively) you need to enable access to Emdros in PHP. The correct way to do this depends on your Apache and PHP installation. In my case, I had to add a file with this contents to the directory `/etc/php5/mods-available`:

```
extension=/usr/local/lib/emdros/EmdrosPHP.so
```

If you cannot or do not want to enhance PHP with Emdros support, you can configure Bible OL to use the *mql* command line tool instead, but this is not recommended. More information about that is given in Section 16.3.3.

6.1.2 Emdros on a Runtime System

On a runtime system, you can install Emdros is described for a development system. Alternatively, if your runtime system uses the same operating system as your development system, you can make and install the Emdros system on your development system and then copy the resulting files to your runtime systems. The files to copy will be typically found in `/usr/local/bin` and `/usr/local/lib/emdros`.

(I have never actually done this, and I suspect that you may need to run *ldconfig* after copying the files to your runtime system.)

TBD: External MQL command execution.

6.2 Installing Lessc

This is only required on a development system.

Lessc is the Less compiler. It runs under *nodejs*, which is a stand-alone JavaScript runtime system. For information about how to install and use the Less compiler, see `http://lesscss.org`. There, they recommend that lessc be installed using this shell command:

```
npm install -g less
```

You may not need the "-g" option. If you do include it, you may need to prefix the command by *sudo*.

I use version 1.3.3 of *lessc*, but the current version on GitHub is 2.4.0. I have not studied what impact, if any, using a new version of the compiler will have on the Less source files in Bible OL.

6.3 Installing Tsc

This is only required on a development system.

Tsc is the TypeScript compiler. It runs under *nodejs*, which is a stand-alone JavaScript runtime system. For information about how to install and use the TypeScript compiler, see `http://www.typescriptlang.org`. There, they recommend that tsc be installed using this shell command:

```
npm install -g typescript
```

You may not need the "-g" option. If you do include it, you may need to prefix the command by *sudo*.

I use version 1.0.1.0 of *tsc*, but the current version on GitHub is 1.4. I have not studied what impact, if any, using a new version of the compiler will have on the TypeScript source files in Bible OL. Currently, the Makefile that comes with Bible OL checks that the compiler version is 1.0.1.0. This is probably unnecessary.

Chapter 7

Source Code Tree Overview

You must read this chapter.

The source code contains the following directories (in alphabetical order):

Name	Contents
`ckeditor`	An HTML text editor. It is used in the server code to allow the user to edit the description of an exercise. Webpage: http://ckeditor.com. Download: https://github.com/ckeditor. License: A choice between GPL, LGPL, and MPL.
`CodeIgniter`	The PHP framework used by the server side of Bible OL. Webpage: http://www.codeigniter.com. Download: https://github.com/bcit-ci/CodeIgniter. License: MIT.
`culmus-fonts`	A collection of Hebrew fonts. These files are not used directly by the server, but a few of the font files from the subdirectory `Squirrel` have been copied to the `styles/fonts` directory. Webpage: http://culmus.sourceforge.net. Download: http://sourceforge.net/projects/culmus/files/culmus/0.130[1] License: GPL.
`db`	The Emdros databases and associated description files.
`images`	Image files used by the server.
`jquery-ui-1.10.2.custom`	A customized version of *jQuery UI*. Used by the client to display components of the user interface. Webpage: http://jqueryui.com. Download: http://jqueryui.com/download. License: MIT.
`js`	JavaScript files from various sources. The files `ol.js`, `editquiz.js`, and `fontselector.js` are the output form compiling TypeScript files. These JavaScript files should therefore never be edited.

(Continued...)

[1] This does not include the `Squirrel` subdirectory. Unfortunately, I don't recall the origin of that directory.

Name	Contents
`jstree`	A JavaScript component used by the server to display a hierarchy of books, chapters, and verses of the Bible. Webpage: http://jstree.com. Download: https://github.com/vakata/jstree. License: A choice between MIT and GPL.
`myapp`	The server code. Chapter 16 gives more information.
`quizzes`	Quizzes available for the user. This directory is used only by the runtime system. Development files should not be stored here, and the contents of the directory is not under Git control.
`quiz_templates`	Sample quiz templates to be copied to the `quizzes` directory in a new installation.
`SILfonts`	A collection of Greek and phonetic fonts. These files are not used directly by the server, but a few of the font files from the sub-subdirectories `*/Squirrel` have been copied to the `styles/fonts` directory. Webpage: http://scripts.sil.org. Download: Search the http://scripts.sil.org website for relevant fonts. License: SIL Open Font License.
`styles`	CSS, Less, and fonts files from various sources.
`techdoc`	The technical documentation.
`ts`	TypeScript source files for the client. See the following entries for information about the subdirectories `jquery` and `jqueryui`.
`ts/jquery`	The TypeScript definitions of the interfaces and variables found in *jQuery*. Webpage: http://definitelytyped.org. Download: https://github.com/borisyankov/DefinitelyTyped/tree/master/jquery. License: MIT.
`ts/jqueryui`	The TypeScript definitions of the interfaces found in *jQuery UI*. Webpage: http://definitelytyped.org. Download: https://github.com/borisyankov/DefinitelyTyped/tree/master/jqueryui. License: MIT.
`uguide`	TBD
`valums-file-uploader-b3b20b1`	An old version of a file upload mechanism, used to upload exercise files to the server. This code was released under a GPL license. Since this code was copied to Bible OL, it ownership and licensing has changed. It is now known as *FineUploader* and is available from these sources: Webpage: http://fineuploader.com. Download: https://github.com/FineUploader/fine-uploader License: Widen Commercial License.[2] (I cannot tell this from their license, but according to their website the license allows royalty-free use for non-commercial purposes.)
`virtualkeyboard`	A JavaScript-based virtual keyboard for typing Greek and Hebrew in the client. Website: http://allanguages.info. Download: http://freecode.com/projects/jsvk. License: LGPL.

(Continued...)

[2] https://github.com/FineUploader/fine-uploader/blob/master/LICENSE

Name	Contents
`zocial`	Icon and styles for setting up a Google login button. Website: http://zocial.smcllns.com. Download: https://github.com/samcollins/css-social-buttons. License: MIT.

Chapter 8

Emdros Databases in Bible OL

Read this chapter if you are going to work with code that accesses the Emdros databases on the server or displays text and exercises in the client.

Bible OL currently supports two text databases:

- ETCBC4, which contains the Hebrew and Aramaic version of the Old Testament.
- nestle1904, which contains the Greek version of the New Testament.

These two databases are described in detail in appendices A and C. This chapter gives only the most important information. One way to learn more about them is to look at the MQL code used for generating these databases. The MQL code is found in the files TBD and TBD.

Previous versions of Bible OL have used two other databases: *WIVU* for the Old Testament and *tisch* for the New Testament. The *WIVU* database was protected by a more restrictive copyright than ETCBC4, and the *tisch* database was based on Tischendorf's Greek New Testament, which used peculiar spellings in a number of places. However, traces of these databases can still be found in the system as described in Section 9.2.

In the Bible OL source tree, the Emdros databases are found in the directory `db`.

8.1 The *visual* Feature

The Emdros databases use various feature names to describe the actual text of the corpus. In ETCBC4, the name of the feature is *text_utf8* when the Hebrew alphabet is used and *text_translit* when a transliterated alphabet is used; in nestle1904, the name of the feature is *surface*.

In order to establish a uniform way to reference this important feature, the Bible OL server and client code uses the name *visual* as an alias for the text feature of the current Emdros database.

8.2 ETCBC4

This section describes a number of features of the ETCBC4 Hebrew/Aramaic database. A more detailed description can be found in Appendix A.

Text in the database come in three different alphabets:

- Hebrew/Aramaic characters encoded in UTF-8. (In the following text, this will be known as the *native* alphabet.)
- Latin transliteration of the text, encoded in UTF-8. (In the following text, this will be known as the *transliterated* alphabet.)
- Hebrew/Aramaic characters int *ETCBC4 transcription*. This transcription is defined in the document `ETCBC4-transcription.pdf` which is located together with the current document or can be downloaded from `http://shebanq.ancient-data.org/shebanq/static/docs/`

ETCBC4-transcription.pdf. (In the following text, this will be known as the *transcribed* alphabet.)

For example, using these three encodings, the three different encodings of the word "created" from Genesis 1 : 1 is encoded as:

- *Native:* בָּרָ֣א
- *Transliterated:* bārāʾ
- *Transcribed:* B.@R@74>

The transcribed characters should never be displayed to users, but they can be useful for internal use because they only use a limited set of ASCII characters.

8.2.1 Object Types

The ETCBC4 database contains these object types:

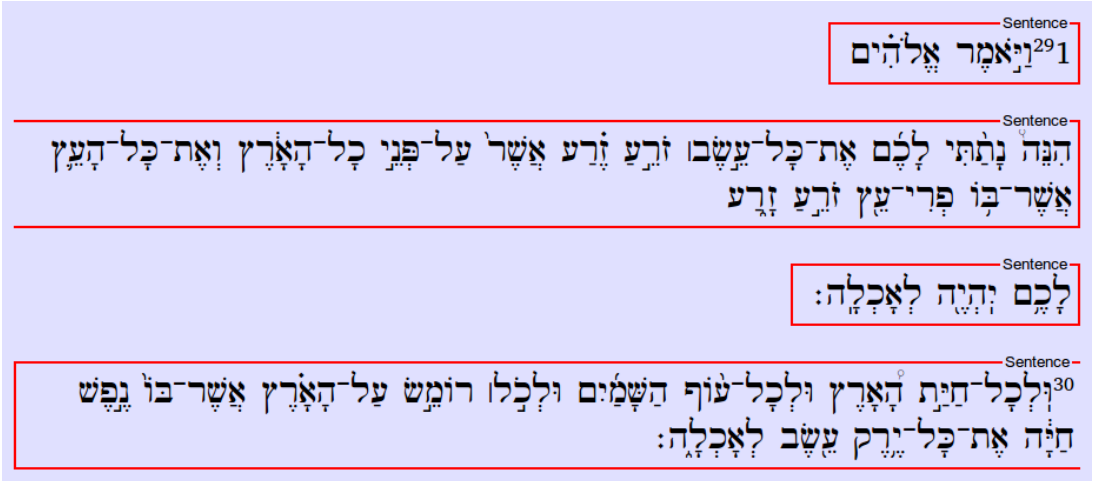
- word
- sentence
- sentence_atom
- clause
- clause_atom
- subphrase
- phrase
- phrase_atom
- book
- chapter
- verse
- half_verse

8.2.1.1 Syntactic Object Types

The object types *word, sentence, sentence_atom, clause, clause_atom, subphrase, phrase,* and *phrase_atom* describe the syntactic composition of the text.

The basic object type is the *word.* Each *word* corresponds to a single monad.

The top-level syntactic element is the *sentence.* A sentence may be built from sets of noncontiguous monads. Each contiguous part of a sentence is a *sentence_atom* object. Consider, for example, Genesis 1 : 29-30:



The sentence starting with the word הִנֵּה consists of two parts. One sentence_atom starts at הִנֵּה and ends at זָרַע, another sentence_atom starts at וּלְכָל־חַיַּת and ends at לְאָכְלָה. Together, these two sentence_atoms make up a single sentence. If a sentence is contiguous, it contains a single sentence_atom.

As illustrated above, Bible OL displays noncontiguous sentences using boxes where one of the sides is missing.

Sentence objects consist of *clause* objects, which – like sentences – are comprised of *clause_atom* objects.

Clause objects consists of *phrase* objects, which are comprised of *phrase_atom* objects.

Phrase objects may contain *subphrase* objects. Note the word "may"; not all words belong to a subphrase. Subphrase objects are always built from contiguous monads. Subphrases may contain other subphrases; for example, in Genesis 1 : 16 the words וַיַּעַשׂ אֱלֹהִים אֶת־שְׁנֵי הַמְּאֹרֹת הַגְּדֹלִים contain three subphrases:

- שְׁנֵי
- הַמְּאֹרֹת הַגְּדֹלִים
- הַגְּדֹלִים

Note how the third subphrase is part of the second subphrase.

8.2.1.2 Editorial Object Types

The object types *book, chapter, verse,* and *half_verse* describe the editorial composition of the text.

The objects *book, chapter,* and *verse* are self-explanatory. The *half_verse* objects identify a subdivsion of verses into two halves, labelled A and B. For example, in Genesis 1 : 1, the two half_verse objects correspond in English to:

A: In the beginning God created
B: the heavens and the earth.

8.2.2 What Is a Word?

In most western languages, a space is inserted between two words. In Hebrew, some words are strung together as one. For example, the text וַיְהִי־אוֹר ("and there was light") in Genesis 1 : 3 consists of the three words וַ ("and"), יְהִי ("there was"), and אוֹר ("light").

In ETCBC4 this issue is handled by associating an Emdros feature called *continuation* with each word. The continuation feature contains

- a space if a space should be inserted between this word and the next,
- an empty string if this word should be strung together with the following word,
- a ־ (Unicode value 05BE) if a *maqaf* (hyphen) should to be inserted between this word and the next.

So for the text וַיְהִי־אוֹר, mentioned above, we have these features for the three words:

Text	Continuation
וַ	Empty string
יְהִי	־
אוֹר	Space

8.3 Nestle1904

This section describes a number of features of the nestle1904 Greek database. A more detailed description can be found in Appendix C.

8.3.1 Object Types

The nestle1904 database contains these object types:

- word
- sentence
- clause1
- clause2
- book
- chapter
- verse

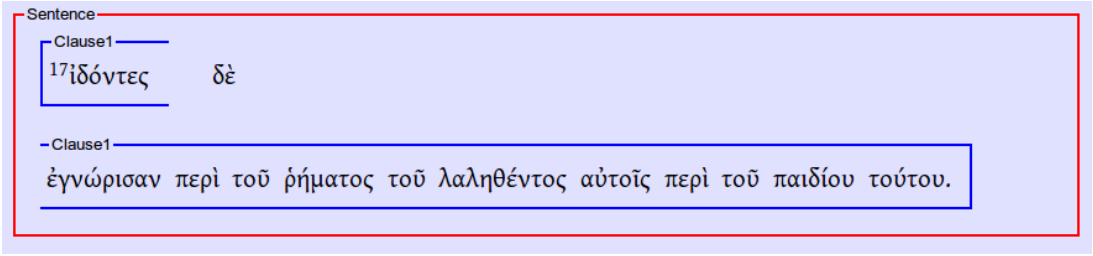
8.3.1.1 Syntactic Object Types

The object types *word, sentence, clause1,* and *clause2* describe the syntactic composition of the text.

The basic object type is the *word.* Each *word* corresponds to a single monad. In contrast to the Hebrew database, the Greek database has no concept of a *continuation* feature (see Section 8.2.2).

The top-level syntactic element is the *sentence.* A sentence is a set of contiguous monads.

Sentence objects contain *clause1* objects. A *clause1* object may be built from sets of noncontiguous monads. Consider, for example, Luke 2 : 17:



This sentence contains a single *clause1* object, but the word *δὲ* is not considered part of that object. As illustrated above, Bible OL displays noncontiguous *clause1* objects using boxes where one of the sides is missing.

Clause1 objects may contain of *clause2* objects, which – like *clause1* objects – may be noncontiguous.

8.3.1.2 Editorial Object Types

The object types *book, chapter,* and *verse* describe the editorial composition of the text.

8.4 A Note on Greek Accents in Unicode

Classical Greek used polytonic accents over vowels. For example, the letter alpha could have these polytonic accents:

Character	Greek accent name	English accent name
ά	Oxia (οξεία)	Acute
ὰ	Varia (βαρεία)	Grave
ᾶ	Perispomeni (περισπωμένη)	Circumflex

Modern Greek is not a polytonic language, and in 1982 the polytonic accents were replaced by a single, monotonic accent: the *tonos* (τόνος). In the early years of the monotonic system, particularly when reformers wished to differentiate their system from the polytonic, the tonos on letters was a novel

sign: typically a dot or wedge: ὰ. However, the Greek government decreed in 1986 that the tonos shall be the acute. So you must now write ά instead of ὰ.[1]

This confusion has had an impact on the representation of the oxia and tonos accents in Unicode. Because of the decision from 1982, Unicode distinguishes between the oxia and the tonos; but because of the decision from 1986, a change was made in Unicode version 3.0 stating that the character ά should be encoded with the Unicode value for GREEK SMALL LETTER ALPHA WITH TONOS, even when writing ancient Greek.

The affected Unicode characters are:

	Unicode value with	
Character	**tonos**	**oxia**
ά	03AC	1F71
έ	03AD	1F73
ή	03AE	1F75
ί	03AF	1F77
ό	03CC	1F79
ύ	03CD	1F7B
ώ	03CE	1F7D
ΐ	0390	1FD3
ΰ	03B0	1FE3

So nowadays the correct way to encode the character ά is to use the value 03AC, regardless of whether the accent is an oxia or a tonos.

But here's the catch: *The nestle1904 database uses the oxia encodings, not the recommended tonos encodings.* (The reason is probably a desire to emphasize that polytonic accents are used.)

This necessitates a conversion between tonos encoding and oxia encoding in a few places in Bible OL.

[1] Source: http://www.tlg.uci.edu/~opoudjis/unicode/unicode_gkbkgd.html, accessed 20 February 2015.

Chapter 9

Database Description Files

Read this chapter if you are going to work with code that accesses the Emdros databases on the server or displays text and exercises in the client.

Each database is associated with a number of files that describe details about the database. This chapter describes these files in detail.

In the Bible OL source tree, the database description files are found in the directory `db`.

In the following sections the file names all start with the string "XXX". The actual file name can be anything, but examples will be given in the text.

Most of the files are JSON files. A JSON file contains key/value pairs, where the value can be a string, a number, a Boolean value, an array of values, or another collection of key/value pairs.

A JSON file can either be "ugly" or "pretty". This is an example of an ugly JSON file:

```
{"alpha":8,"beta":{"gamma":true,"delta":["ten","eleven"]}}
```

This is the same data in pretty format:

```
{
    "alpha": 8,
    "beta": {
        "gamma": true,
        "delta": [
            "ten",
            "eleven"
        ]
    }
}
```

Both of these listings describe the same object. The object contains two key/value pairs:

- "alpha" with the numerical value 8.
- "beta" with a value that is a collection of key/value pairs.

The key "beta" has a value that contains two key/value pairs:

- "gamma" with the Boolean value *true*.
- "delta" with a value that is an array containing the two strings "ten" and "eleven".

Bible OL works equally well with ugly and pretty JSON files, but the ugly format is normally preferred because it takes up less space (and makes reverse engineering slightly more difficult for the uninitiated). The pretty format is, of course, easier for humans to understand is therefore useful while debugging the system.

The script *json_pretty_print.php* can be used to convert between the ugly and the pretty format. If the file "xxx.json" contains JSON data (either ugly or pretty), the command

```
php json_pretty_print.php -p xxx.json
```

will write the data in pretty format to standard output; and the command

```
php json_pretty_print.php -u xxx.json
```

will write the data in ugly format to standard output.

Note that *json_pretty_print.php* is completely unforgiving about errors in its input. If the input JSON file contains the slightest error (such as a missing or extra comma), the script will simply output the word "null".

9.1 Database Specification File: XXX.db.json

On the server the *Database Specification File* has a name that ends in `.db.json`. For Bible OL, this is the main point of access to the Emdros databases. When Bible OL needs to list the available databases, it searches for files with names that end in `.db.json`.

On the client the contents of the Database Specification File is available in a variable called *configuration*. Its structure is described in TypeScript as the *Configuration* interface in the file `ts/configuration.ts`.

The Database Specification File describes how the individual parts of the database are used by Bible OL. It describes what features are available for exercises and what grammatical features the user can choose to display.

Multiple Database Specification Files can refer to the same Emdros database. For example, `ETCBC4.db.json` and `ETCBC4-translit.db.json` both reference the ETCBC4 Emdros database, but the former displays text using the native alphabet, whereas the latter displays text using the transliterated alphabet.

The Database Specification File is a JSON file containing the following key/value pairs:

Key	Value
version	A number which identifies the layout used by this file. This number is currently ignored.
databaseName	The name of the Emdros database file. Currently, this is either "ETCBC4" or "nestle1904".
propertiesName	The name of the Database Localization Files (see Section 9.3).
databaseVersion	A string containing the version number of the database. This number is only used for display purposes. Whenever an Emdros database is changed, this number should also be changed.
granularity	The name of the Emdros object type defining the amount of text to display in an exercise. Typically, this name is "sentence".
surfaceFeature	The name of the feature that contains the actual text to display. For ETCBC4 using the native alphabet this value is "text_utf8", for ETCBC4 using the transliterated alphabet the value is "text_translit", for nestle1904 the value is "surface".
objHasSurface	The name of the Emdros object type that contains the *surfaceFeature*.
suffixFeature	The name of the feature that contains the continuation for the word to display (see Section 8.2.2). For ETCBC4 using the native alphabet this value is "continuation_utf8", for ETCBC4 using the transliterated alphabet the value is "continuation_translit", for nestle1904 the value is *null*.

(Continued...)

Key	Value
charSet	The name of the character set for the text. For ETCBC4 using the native alphabet this value is “hebrew”, for ETCBC4 using the transliterated alphabet the value is “transliterated_hebrew”, for nestle1904 the value is “greek”.
objectSettings	A collection of key/value pairs containing information about how Bible OL should treat Emdros object type. This is detailed in Section 9.1.1.

universeHierarchy	An array containing information about how the text references are structured. For the Bible, this hierarchy is book/chapter/verse. A typical value is
	`"universeHierarchy": [`
	`{`
	`"type": "book",`
	`"feat": "book"`
	`},`
	`{`
	`"type": "chapter",`
	`"feat": "chapter"`
	`},`
	`{`
	`"type": "verse",`
	`"feat": "verse"`
	`}`
	`]`
	This means that the top reference level is found in the *book* feature of the Emdros type *book*, the second reference level is found in the *chapter* feature of the Emdros type *chapter*, and the third reference level is found in the *verse* feature of the Emdros type *verse*.

picDb	The URL of the resource website (see Section 20.2). This value may be *null.*
sentencegrammar	An array containing information about the grammar items available to the user. This is detailed in Section 9.1.3.
subsetOf	If the Database Specification File describes a subset of a larger database, *subsetOf* gives information about the subset. This is detailed in Section 9.1.7. This value is always *null* for the ETCBC4 and nestle1904 databases.
useSofPasuq	A Boolean value indicating if a sof pasuq (:) should be added to the end of a Hebrew verse. This key is only present for Hebrew text.

9.1.1 The *objectSettings* Key

The *objectSettings* key in the Database Specification File gives detailed information about how the Emdros object types and their features should be treated by Bible OL. The *objectSettings* key has a value that is a collection of key/value pairs, where the keys are Emdros object type. The corresponding values give information about how the Emdros object should be treated.

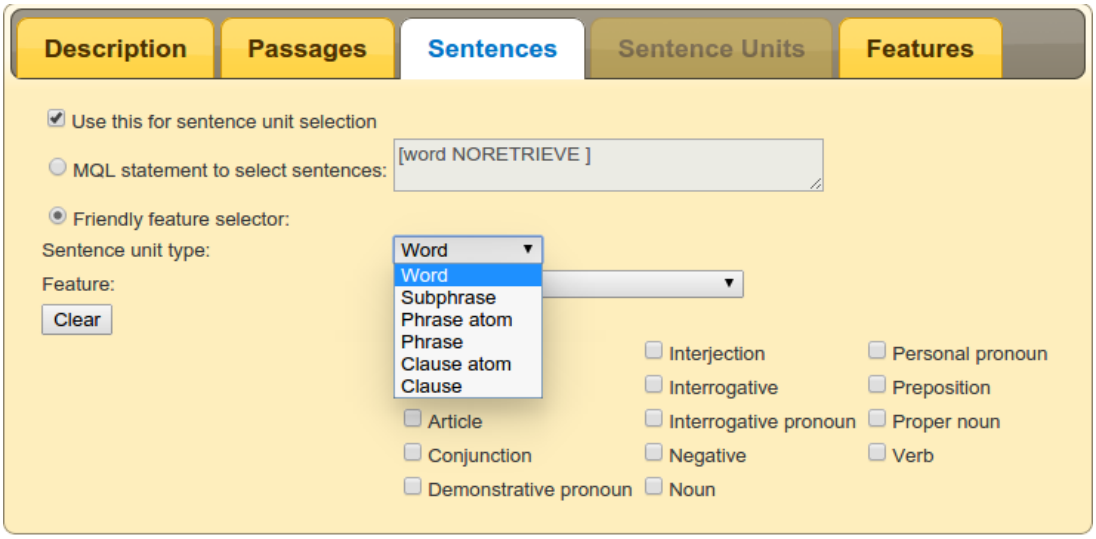
Let us consider a subset of the *objectSettings* for the ETCBC4 database:

```
"objectSettings": {
    "book": {
```

```
        },
        "chapter": {
        },
        "verse": {
        },
        "word": {
            "mayselect": true,
            "additionalfeatures": [...],
            "featuresetting": {...}
        },
        "subphrase": {
            "mayselect": true,
            "featuresetting": {...}
        }
    }
```

In this subset, the Emdros types *book, chapter, verse, word,* and *subphrase* are mentioned. No special information about *book, chapter,* and *verse* is provided, which means that these Emdros types cannot be made the subject of exercises. It is the presence of the key *mayselect* with the value *true* that signals to Bible OL that the associated Emdros object can be used when select quiz objects for an exercise. So in the above example, the Emdros objects *word* and *subphrase* can be used for quiz object selection.

When you are creating an exercise with the ETCBC4 database, the "Sentences" and "Sentence Units" tab allow you to specify a sentence unit type:



The values in the drop down box are the Emdros object that have *mayselect* set to *true* in *objectSettings.* As shown in the example above, these objects have one or two additional key/value pairs with the keys *featuresetting* and, optionally, *additionalfeatures.*

The value of *featuresetting* is another collection of key/value pairs, giving details about the features of the object. This is detailed in Section 9.1.2.

The *additionalfeatures* key identifies an array of features that should always be retrieved for this Emdros object, even though these features are not the subject of an exercise. They are used for accessing the multiple choice database (see Chapter 11).

9.1.2 The *featuresetting* Key

The *featuresetting* key under an Emdros object type gives detailed information about how the features of the Emdros object should be treated by Bible OL. Its value is a collection of key/value pairs, where the keys are feature names. The corresponding values give information about how the feature should be treated.

Let us consider a subset of the *featuresetting* for the *word* object in the ETCBC4 database:

```
"featuresetting": {
    "lexeme_occurrences": {
        "ignoreShowRequest": true,
        "isRange": true
    },
    "english": {
        "ignoreSelect": true,
        "matchregexp": "\/^(.+[;,] +)?{0}([,;].+)?$\/i"
    },
    "text_cons_utf8": {
        "hideWord": true,
        "foreignText": true
    },
    "text_nocant_utf8": {
        "alternateshowrequestDb": "ETCBC4_words.db",
        "alternateshowrequestSql": "SELECT DISTINCT word FROM texts,lextext,lexemes WHERE
lex='%s' AND lexid=lexemes.id AND textid=texts.id",
        "hideWord": true,
        "foreignText": true
    },
    "vt": {
        "hideValues": [
            "weyq"
        ]
    }
}
```

In this subset, the features *lexeme_occurrences, english, text_const_utf8, text_nocant_utf8,* and *vt* are mentioned. The value associated with each of these keys is another collection of key/value pairs. Many of the values are Boolean, and an absent key is equivalent to a Boolean value of *false*. Thus, the absence of a *hideWord* key from *lexeme_occurrences* has the same meaning as if *hideWord* had been given the value *false*.

The following table lists the keys and values that can be associated with a feature of an Emdros object:

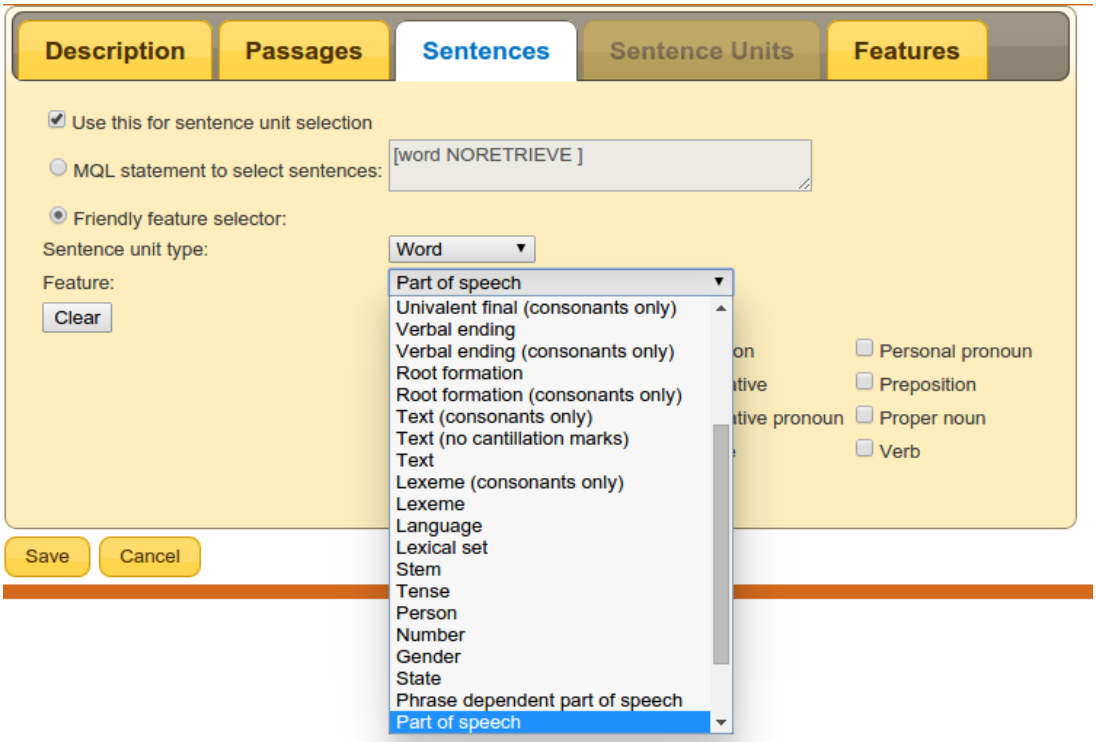
Key	Value
ignoreSelect	A value of *true* means: Do not use this feature for object selection (see below).
isDefault	This value must be *true* for exactly one feature of an Emdros object. It indicates that this feature is the initially displayed feature in the "Sentences" tab when creating an exercise (see below).
ignoreShow	A value of *true* means that this feature cannot be displayed as part of an exercise. In other words, there is no "Show" button for this feature on the "Features" tab when creating an exercise.
ignoreRequest	A value of *true* means that this feature cannot be requested as part of an exercise. In other words, there is no "Request" button for this feature on the "Features" tab when creating an exercise.
ignoreShowRequest	A value of *true* means that both *ignoreShow* and *ignoreRequest* should be taken as *true*.[1]

(Continued...)

[1] The intention is to remove this key from *featuresetting* in the future.

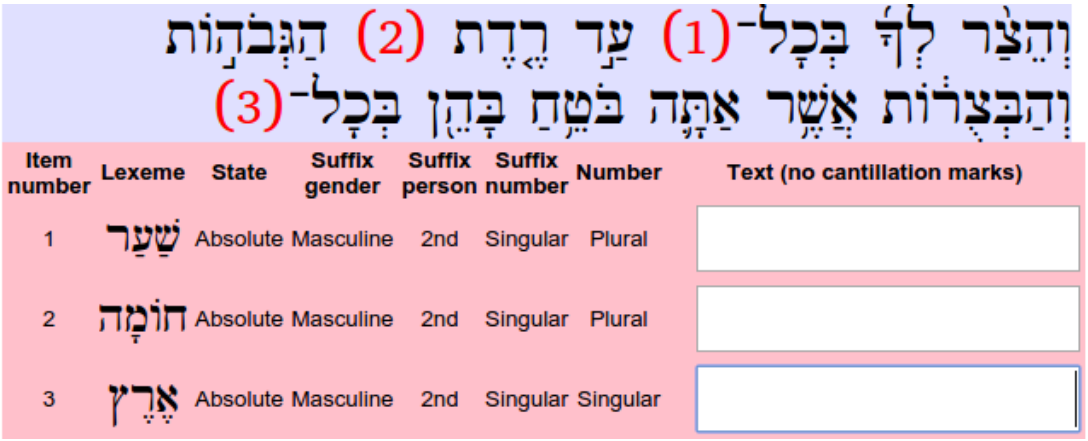
Key	Value
hideWord	If this value is *true* and the feature is a request feature for an exercise, the corresponding words should be replaced by a number in the displayed text (see below).
foreignText	A value of *true* means that this feature is written using a non-Latin alphabet.
transliteratedText	A value of *true* means that this feature is written using the transliterated Hebrew alphabet.
hideValues	Relevant only for enumeration features. It is an array of enumeration values that never occur in a text and should therefore be omitted from the user interface.
isRange	A value of *true* means that this feature represents range of integer values.
otherValues	An array of enumeration feature values that should be lumped together as "Other" in the user interface.
matchregexp	A regular expression used to check if an answer provided by a learner matches the value of a feature. For example, the *english* feature for the Hebrew word אֶרֶץ has the value "land; territory, country; the earth". The regular expression in *matchregexp* is designed to ensure that a learner's answer is accepted, regardless of whether the answer is "land", "territory", "country", or "the earth".
alternateshowrequestDb	The name of a multiple choice database (see Chapter 11).
alternateshowrequestSql	An SQL statement used to access the multiple choice database (see Chapter 11).

The keys *ignoreSelect* and *isDefault* control the contents of the feature selection menu when creating an exercise:



If *ignoreSelect* is set for a feature, then that feature is not shown in the selection menu. The feature with *isDefault* set is the selected feature when the dialog is first loaded.

If *hideWord* is *true* and the feature is used as a request feature in an exercise, the corresponding word is replaced by a number in the text. For example, in the following exercise the feature *text_nocant_utf8* (that is, "Text (no cantillation marks)") is used as a request feature. Consequently the corresponding words in the text are replaced by the numbers (1), (2), and (3) lest the words in the text give the answer to the questions:

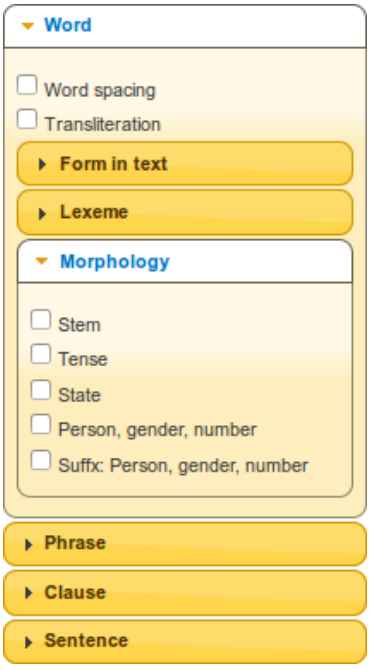


9.1.3 The *sentencegrammar* Key

The *sentencegrammar* key in the Database Specification File gives detailed information about the grammar items available to the user. Its value is an array, in which each entry corresponds to an Emdros object type.

Bible OL uses the information in *sentencegrammar* in two locations. One is the grammar selection box in the left part of the screen, the other is in the grammar information box in the right part of the screen.

The grammar selection box may look like this:



In the above illustration, each of the four main boxes, *Word, Phrase, Clause,* and *Sentence,* of the grammar selection box corresponds to an entry at the top level of *sentencegrammar*.

The grammar information box may look like this:

Word	
Text	שָׁמַ֥יִם
Transliteration	šš̄āmayim
Form in text:	
Root formation	-
Preformative	-
Verbal ending	-
Nominal ending	ם֥
Pronominal suffix	-
Univalent final	-
Qere	-
Lexeme:	
Lexeme	שָׁמַ֥יִם
Lexeme (transliterated)	šāmayim
English	heaven
German	Himmel
Occurrences	421
Frequency rank	123
Part of speech	Noun
Phrase dependent part of speech	Noun
Lexical set	None
Verb class	N/A
Morphology:	
Stem	None
Tense	None
State	Absolute
Person, gender, number	-MPI
Suffx: Person, gender, number	---

In this illustration, the *sentencegrammar* for the *word* object has been used to determine what information to retrieve.

Let us consider the *sentencegrammar* for the ETCBC4 database (taken from the file `db/ETCBC4.db.json`):

```
"sentencegrammar": [
    {
        "mytype": "SentenceGrammar",
        "objType": "word",
        "items": [...]
    },
    {
        "mytype": "SentenceGrammar",
        "objType": "phrase",
        "items": [...]
    },
    {
        "mytype": "SentenceGrammar",
        "objType": "clause",
        "items": [...]
```

```
    },
    {
        "mytype": "SentenceGrammar",
        "objType": "sentence"
    }
]
```

As this example shows, *sentencegrammar* is an array whose elements have three key/value pairs:

- *mytype* with a value that is always "SentenceGrammar".
- *objType* with a value that identifies an Emdros object.
- *items* (optional), which is an array of *GrammarFeature, GrammarMetaFeature,* or *GrammarGroup* specifications, as detailed below.

If you compare the listing of *sentencegrammar* above, with the grammar selection box shown on page 34, you will notice that each element in the *sentencegrammar* corresponds to a top level box in the illustration.

For each Emdros object (that is, for each element in the *sentencegrammar* array) the value of *items* is an array that specified the items that can be displayed for the particular Emdros object. Each entry in the *items* array can have one of three forms: *GrammarFeature, GrammarMetaFeature,* or *GrammarGroup*, as specified in their *mytype* value.

A typical *item* array may look like this:

```
"items": [
    {
        "mytype": "GrammarFeature",
        "name": "text_translit"
    },
    {
        "mytype": "GrammarGroup",
        "name": "form_in_text",
        "items": [...]
    },
    {
        "mytype": "GrammarMetaFeature",
        "name": "pgn",
        "items": [...]
    }
]
```

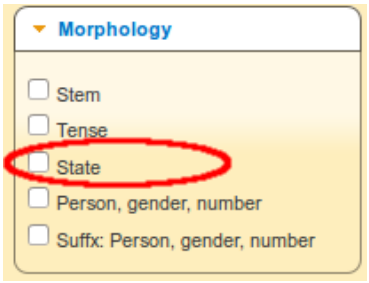
In this example, the *items* array has two elements, one with *mytype=GrammarFeature*, one with *mytype=GrammarGroup*, and one with *mytype=GrammarMetaFeature.*

9.1.4 GrammarFeature

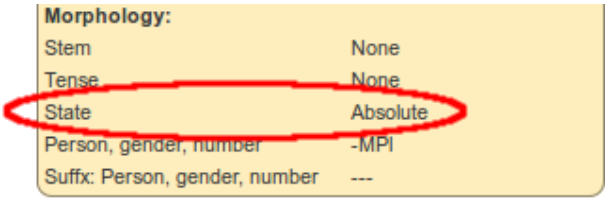
A *sentencegrammar* item with *mytype=GrammarFeature* describes a single Emdros feature. The format is:

```
{
    "mytype": "GrammarFeature",
    "name": an Emdros feature name
}
```

In the grammar selection box, a GrammarFeature is displayed thus:



In the grammar information box, a GrammarFeature is displayed thus:



9.1.5 GrammarMetaFeature

A *sentencegrammar* item with *mytype=GrammarFeature* describes a combined value of a number of Emdros features.

A GrammarMetaFeature is specified thus:

```
{
    "mytype": "GrammarMetaFeature",
    "name": the name of the GrammarMetaFeature,
    "items": [
        {
            "mytype": "GrammarSubFeature",
            "name": an Emdros feature name
        },
        {
            "mytype": "GrammarSubFeature",
            "name": an Emdros feature name
        },
        ... Additional GrammarSubFeatures
    ]
}
```

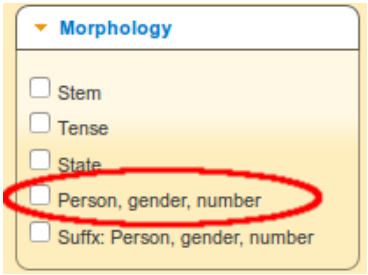
In the *items* array the Emdros features that make up the GrammarMetaFeature are listed with a *mytype* value of "GrammarSubFeature".

As an example, Bible OL combines the person, gender, and number of a word to a single item, such as "2FSg" (which means 2nd person, feminine, singular). This is specified thus:

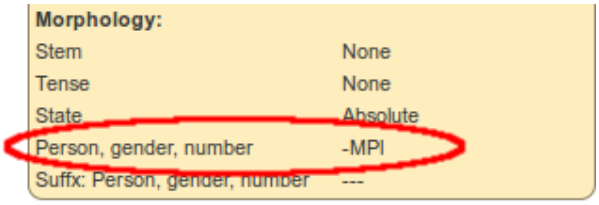
```
{
    "mytype": "GrammarMetaFeature",
    "name": "pgn",
    "items": [
        {
            "mytype": "GrammarSubFeature",
            "name": "ps"
        },
        {
            "mytype": "GrammarSubFeature",
            "name": "gn"
```

```
            },
            {
                "mytype": "GrammarSubFeature",
                "name": "nu"
            }
        ]
    },
```

The *ps, gn,* and *nu* features represent the person, gender, and number of a word, respectively.
In the grammar selection box, a GrammarMetaFeature is displayed thus:



In the grammar information box, a GrammarMetaFeature is displayed thus:



9.1.6 GrammarGroup

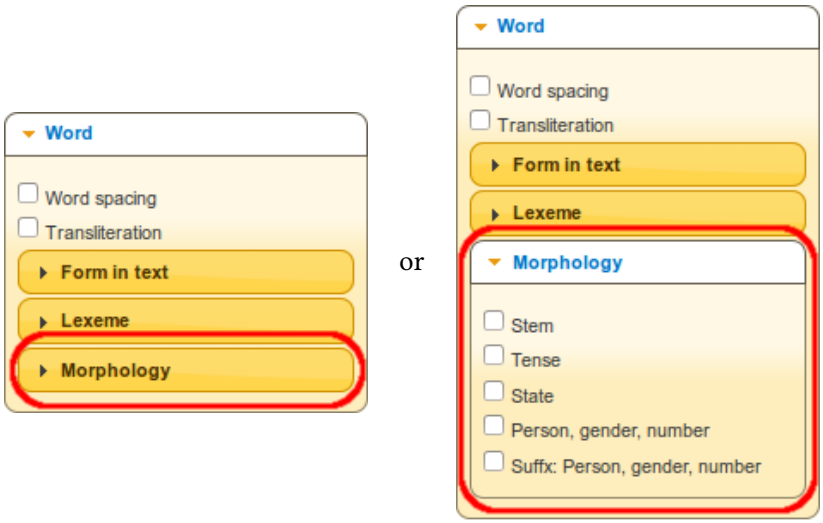
A *sentencegrammar* item with *mytype=GrammarGroup* groups GrammarFeatures and GrammarMetaFeatures into logical units, such as "Features that describe the lexeme" or "Features that describe the morphology".

A GrammarGroup is specified thus:

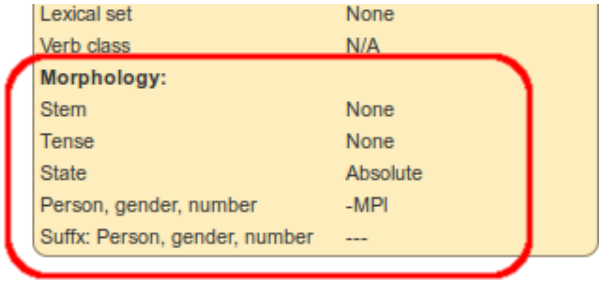
```
{
    "mytype": "GrammarGroup",
    "name": the name of the GrammarGroup,
    "items": [...]
}
```

The *items* array here contains a collection of GrammarFeatures and GrammarMetaFeatures in the format described above.

In the grammar selection box, a GrammarGroup (called "Morphology" in this case) is displayed thus:



In the grammar information box, a GrammarGroup is displayed thus:



9.1.7 The *subsetOf* Key

The *subsetOf* value is present for historical reasons. With the current ETCBC4 and nestle1904 databases it is not required; however, the previously used WIVU database, was available in two versions: the entire Old Testament and a subset containing about 20% of the text.

The *subsetOf* key specifies the relationship between the subset and the superset. It is used to enable exercises that have been created for the subset to be used with the superset.

The format of *subsetOf* is illustrated by this example taken from the subset Database Specification File:

```
"subsetOf": {
    "name": "WIVU",
    "properties": "WIVU",
    "provides": [
        "Genesis",
        "Exodus:1",
        "Exodus:2",
        "Exodus:3",
        ... (Additional Bible references omitted)
    ]
}
```

Here, *name* is the name of the Emdros database file for the superset, and *properties* is the name of the Database Localization Files (see Section 9.3) for the superset.

The *provides* array identifies the books and chapters provided by the subset. In this example, the subset includes the entire book of Genesis and the first three chapters of Exodus.

If *subsetOf* has the value *null*, the current database is not a subset of any other database.

9.2 Database Description Files for Old Databases

Currently, Bible OL uses two databases, the Hebrew *ETCBC4* database and the Greek *nestle1904* database. Previous versions of Bible OL used the Hebrew *WIVU* database (and various variants of it) and the Greek *tisch*[2] database.

The statistics tables in the user database (see Section) may still contain data from these older databases. In order to display these statistics, Bible OL needs description and localization information for the databases. For this reason, database descriptions files for the old databases are still present in the `db` directory.

9.3 Database Localization Files: XXX.LANG.prop.json

All keys and values in the Database Specification File (Section 9.1) are language independent. **On the server** the *propertiesName* key (page 29) of the Database Specification File is used to locate language specific files for the database, the so-called *Database Localization Files*. If the *propertiesName* key has a value of ABC, the file `db/ABC.LANG.prop.json` will contain localization information for the language LANG. For English, LANG should be "en".

For example, in the Database Specification File `db/ETCBC4-translit.db.json` the value of *propertiesName* is "ETCBC4-translit". The corresponding English localization information is therefore found in the file file `db/ETCBC4-translit.en.prop.json`.

The language is selected by the server. Note: Currently, only English localization files exist, and the language code "en" is hard-coded into the PHP files.[3]

On the client the contents of the selected Database Localization File is available in a variable called *localization*. Its structure is described in TypeScript as the *Localization* interface in the file `ts/localization.ts`.

The Database Localization File is a JSON file containing the following key/value pairs:

Key	Value
dbdescription	A short description of the database.
dbcopyright	An HTML string containing copyright information for the database.
emdrosobject	A collection of key/value pairs containing the localized names for the Emdros object types and their features (see Section 9.3.1).
emdrostype	A collection of key/value pairs containing the localized names for the values in the Emdros enumeration types (see Section 9.3.2).
grammarfeature	A collection of key/value pairs giving the names of GrammarFeatures (see Section 9.3.3).
grammarmetafeature	A collection of key/value pairs giving the names of GrammarMetaFeatures (see Section 9.3.3).
grammarsubfeature	A collection of key/value pairs giving the names of features within a GrammarMetaFeature (see Section 9.3.3).
grammargroup	A collection of key/value pairs giving the names of GrammarGroups (see Section 9.3.3).
universe	A collection of key/value pairs describing how to display book, chapter, and verse references (see Section 9.3.4).

[2] Short for *Tischendorf*.

[3] In the file `myapp/models/mod_askemdros.php` in the functions *setup()* and *db_and_books()*, and in the file `myapp/controllers/ctrl_statistics.php` in the function *show_stat()*.

9.3.1 The *emdrosobject* Key

The value of *emdrosobject* is a collection of key/value pairs containing the localized names for the Emdros object types and their features. As an example, let us consider a subset of the *emdrosobject* for English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"emdrosobject": {
    "word": {
        "_objname": "Word",
        "vt": "Tense",
        "sp": "Part of speech",
        ... (Additional features omitted)
    },
    "phrase_atom": {
        "_objname": "Phrase atom",
        "det": "Determination",
        "rela": "Relation",
        ... (Additional feature omitted)
    },
    ... (Additional Emdros object types omitted)
}
```

Each key within the *emdrosobject* is the name of an Emdros object, so the above example gives information about the *word* and the *phrase_atom* Emdros objects.

Each key has a value which is a collection of key/value pairs. One of those keys is always *_objname*, and its value is the English name for the Emdros object; the remaining keys are features of the Emdros object, and their values are the English name for the feature.

So the above example states that the English name of the Emdros object *word* is "Word", and that object has a feature called *vt* which in English should be rendered as "Tense". The *word* feature *sp* should be rendered "Part of speech" in English.

Similarly, the English name of the Emdros object *phrase_atom* is "Phrase atom", and that object has a feature called *det* which in English should be rendered as "Determination". The *phrase_atom* feature *rela* should be rendered "Relation" in English.

9.3.2 The *emdrostype* Key

The value of *emdrostype* is a collection of key/value pairs containing the localized names for the value of Emdros enumeration types. As an example, let us consider a subset of the *emdrostype* for English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

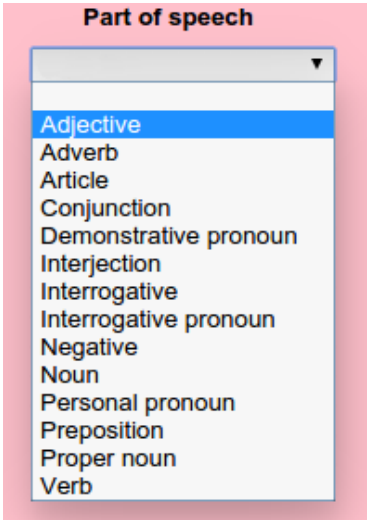
```
"emdrostype": {
    "part_of_speech_t": {
        "verb": "Verb",
        "subs": "Noun",
        "nmpr": "Proper noun",
        ... (Additional values omitted)
    },
    "gender_t": {
        "f": "#2 Feminine",
        "m": "#1 Masculine",
        "NA": "#3 None",
        "unknown": "#4 Unknown"
    },
    ... (Additional enumeration types omitted)
}
```

Each key within the *emdrostype* is the name of an Emdros enumeration type, so the above example gives information about the *part_of_speech_t* and the *gender_t* enumeration types.

Each key has a value which is a collection of key/value pairs, giving the names and the English translation of the values of the enumeration type.

In the above example, the type *part_of_speech_t* has values such as *verb, subs,* and *nmpr*, whose English translations are "Verb", "Noun", and "Proper noun", respectively.

The type *gender_t* has values *f, m, NA,* and *unknown*, whose English translations are "Feminine", "Masculine", "None", and "Unknown", respectively. The strings "#1", "#2" etc. indicate the order in which these values should be sorted when presented to the user. Normally, the values would be sorted alphabetically thus:



But if the English translation starts with "#1", "#2" etc. these numbers indicate the sort order. So with the contents of *gender_t* given above, genders are sorted thus:

9.3.3 The *grammarfeature, grammarmetafeature, grammarsubfeature,* and *grammargroup* Keys

Section 9.1.3 describes how the Database Specification File specifies how grammar information should be grouped in the grammar selection box and the grammar information box on the Bible OL webpage. Sections 9.1.4, 9.1.5, and 9.1.6 describe the GrammarFeature, GrammarMetaFeature, and GrammarGroup specifications and the GrammarSubFeature which is part of a GrammarMetaFeature.

In the Database Localization File, the *grammarfeature, grammarmetafeature, grammargroup,* and *grammarsubfeature* keys give the translation of these items, as detailed below.

9.3.3.1 *grammarfeature*

As an example, let us consider the *grammarfeature* for the English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"grammarfeature": {
    "word": {
        "text_translit": "Transliteration"
    }
}
```

Section 9.3.1 describes how the *emdrosobject* key is used to provide translations for Emdros object features. The *grammarfeature* in the above example gives an alternative translation. Normally the translation of a feature is taken from *emdrosobject*, but in the case of the grammar specification box and the grammar information box, the translation in *grammarfeature* is used, if present. If no translation is given in *grammarfeature*, the translation from *emdrosobject* is used.

9.3.3.2 *grammarmetafeature*

This is the *grammarmetafeature* for the English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"grammarmetafeature": {
    "word": {
        "pgn": "Person, gender, number",
        "suffix_pgn": "Suffix: Person, gender, number"
    }
}
```

Here the GrammarMetaFeature *pgn* of the *word* object is given the English translation "Person, gender, number". (The example in Section 9.1.5 specifies that the *word* object has a GrammarMetaFeature called *pgn*.)

9.3.3.3 *grammarsubfeature*

This is a subset of the *grammarmetafeature* for the English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"grammarsubfeature": {
    "word": {
        "ps": {
            "p1": "1",
            "NA": "-",
            "p2": "2",
            "p3": "3",
            "unknown": "?"
        },
        "gn": {
            "f": "F",
            "m": "M",
            "NA": "-",
            "unknown": "?"
        },
        "nu": {
            "du": "Du",
            "NA": "-",
            "pl": "Pl",
            "sg": "Sg",
            "unknown": "?"
        },
        ... (Additional features omitted)
    }
}
```

The example in Section 9.1.5 specifies that the *word* object has a GrammarMetaFeature called *pgn* which is made up of the GrammarSubFeatures *ps, gn,* and *nu.* The *grammarsubfeature* value listed above specifies the English translation for these three GrammarSubFeatures. So if, for example, a word has features *ps=p2, gn=m,* and *nu=sg* (corresponding to second person, masculine, singular), the *pgn* GrammarMetaFeature should be rendered as "2MSg".

9.3.3.4 *grammargroup*

This is the *grammargroup* for the English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"grammargroup": {
    "word": {
        "form_in_text": "Form in text",
        "lexeme": "Lexeme",
        "morphology": "Morphology"
    }
},
```

The Database Specification File for the ETCBC4 database defines three GrammarGroups for the *word* object with the names *form_in_text, lexeme,* and *morphology.* The localization information above specifies the English names for these GrammarGroups. The two illustrations in Section 9.1.3 show these translations in the grammar selection box and the grammar information box.

9.3.4 The *universe* Key

The value of *universe* is collection of key/value pairs describing how to display book, chapter, and verse references.[4]

As an example, let us consider a subset of the *universe* for English localization of the ETCBC4 database (taken from the file `db/ETCBC4.en.prop.json`):

```
"universe": {
    "book": {
        "_label": "%s",
        "Genesis": "Genesis",
        "Exodus": "Exodus",
        "Leviticus": "Leviticus",
        "Numeri": "Numbers",
        "Deuteronomium": "Deuteronomy",
        "Josua": "Joshua",
        "Judices": "Judges",
        "Ruth": "Ruth",
        "Samuel_I": "1 Samuel",
        "Samuel_II": "2 Samuel",
        ... (Additional books omitted)
    },
    "chapter": {
        "_label": "Chapter %s"
    },
    "verse": {
        "_label": "Verse %s"
    }
}
```

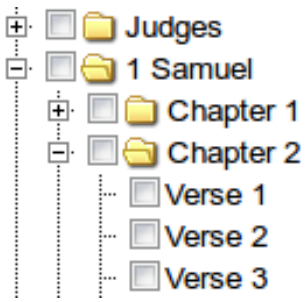
[4]Strictly speaking, references do not have to be Bible references, since Bible OL can handle any other text corpuses. But book, chapter, and verse are used here as a useful illustration.

The *universeHierarchy* key of the Database Specification File (see page 30) defines the reference hierarchy of ETCBC4 as consisting of the Emdros object types *book*, *chapter*, and *verse*. The *universe* key in the above localization example, define how these three object types should be rendered:

The *_label* key presents the general format as a string where "%s" is to be replaced by the actual reference. So, when the *_label* key of *chapter* is "Chapter %s", it means that chapter 18 will be displayed as "Chapter 18".

For the *book* object, the *_label* key is simply the string "%s", but additionally English translations of the book names used in ETCBC4 are given.

Currently, this *universe* information is only used when displaying a passage tree in Bible OL:



9.4 Database Type Information Files: XXX.typeinfo.json

The Emdros databases in Bible OL have an associated Database Type Information File. This file contains information about the Emdros object types and enumeration types. **On the server**, if the name of the Emdros database is XXX, the name of the associated Database Type Information File is `XXX.typeinfo.json`.

On the client the contents of the Database Type Information File is available in a variable called *typeinfo*. Its structure is described in TypeScript as the *TypeInfo* interface in the file `ts/configuration.ts`.

The Database Type Information file is a JSON file. Its contents can be automatically generated from the Emdros database itself. The PHP script `myapp/controllers/ctrl_maketypeinfo.php` contains code to do that. Running the command

```
php index.php maketypeinfo index databasename
```

from the base of the source code tree will generate the type information form the specified Emdros database and write it to standard out in ugly JSON format.

The Database Localization File contains the following key/value pairs:

Key	Value
objTypes	An array containing the names of all Emdros object types.
obj2feat	A collection of key/value pairs that list the features and type of each Emdros object type (see Section 9.4.1).
enymTypes	An array containing the names of all enumeration types in the database.
enum2values	A collection of key/value pairs that list the values of all enumeration types (see Section 9.4.2).

9.4.1 The *obj2feat* Key

The value of the *obj2feat* key is a collection of key/value pairs that list the features and type of each Emdros object type.

This is a subset of *obj2feat* for the ETCBC4 database (taken from the file `db/ETCBC4.typeinfo.json`):

```
"obj2feat": {
    "word": {
        "frequency_rank": "integer",
        "continuation": "string",
        "nu": "number_t",
        "gn": "gender_t",
        "sp": "part_of_speech_t",
        "verb_class": "list of verb_class_t"
        ... (Additional features omitted)
    },
    "clause_atom": {
        "code": "integer",
        "dist": "integer",
        "is_root": "boolean_t",
        "typ": "clause_atom_type_t"
        ... (Additional features omitted)
    },
    ... (Additional object types omitted)
}
```

This example shows that the *word* object has a feature called *frequency_rank* of type *integer*.

9.4.2 The *enum2values* Key

The value of the *enum2values* key is a collection of key/value pairs that list the values of all enumeration types.

This is a subset of *enum2values* for the ETCBC4 database (taken from the file `db/ETCBC4.typeinfo.json`):

```
"enum2values": {
    "boolean_t": [
        "false",
        "true"
    ],
    "number_t": [
        "NA",
        "du",
        "pl",
        "sg",
        "unknown"
    ],
    ... (Additional enumeration types omitted)
}
```

This example shows that the enumeration type *boolean_t* has the values *false* and *true*.

9.5 Database Book Order Files: XXX.bookorder

The Emdros databases in Bible OL have an associated Database Book Order File. If the name of the Emdros database is XXX, the name of the associated Database Book Order File is `XXX.bookorder`. This information is only available **on the server**.

This is a text file that lists the names of the books in the database in the order in which they should be presented. I also lists the chapters available in each book.

This is a subset of the Database Book Order File for the ETCBC4 database (taken from the file `db/ETCBC4.bookorder`):

```
Genesis/1-50
Exodus/1-40
Leviticus/1-27
Numeri/1-36
Deuteronomium/1-34
Josua/1-24
Judices/1-21
Samuel_I/1-31
Samuel_II/1-24
... (Additional books omitted)
```

This file defines the Hebrew order of the books of the Old Testament.[5] Each line consists of the name of the book (as it is defined in the database), followed by a slash and the list of available chapters.

For the ETCBC4 and nestle1904 databases, the chapters are always given as a simple range, such as "1-50" for Genesis, but as mentioned in Section 9.1.7, it is possible to define a subset of a database. In that case a line of the Database Book Order File may look like this:

```
Leviticus/2,6-9,23
```

This line indicates that only chapters 2, 6-9, and 23 of Leviticus are available.

[5] This differs from the order of books used in Christian Bibles.

Chapter 10

Quiz Templates

Read this chapter if you need to understand how exercises are stored in the server.

A quiz template (or an exercise template) is an XML file that describes how Bible OL should generate an exercise. It always has a filename that ends in `.3et`.

A typical quiz template may look like this:

Listing 10.1: Quiz template sample

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <questiontemplate version="3">
3    <desc><![CDATA[Which gender is this?]]></desc>
4    <database>ETCBC4</database>
5    <properties>ETCBC4</properties>
6    <path>Genesis:1:1</path>
7    <path>Genesis:1:2</path>
8    <path>Genesis:3</path>
9    <path>Exodus</path>
10   <sentenceselection version="1">
11     <questionobject>word</questionobject>
12     <featurehandlers version="2">
13       <enumfeature version="1">
14         <name>sp</name>
15         <comparator>equals</comparator>
16         <value>subs</value>
17         <value>prps</value>
18       </enumfeature>
19       <enumfeature version="1">
20         <name>gn</name>
21         <comparator>differs</comparator>
22         <value>NA</value>
23         <value>unknown</value>
24       </enumfeature>
25     </featurehandlers>
26     <useforquizobjects>true</useforquizobjects>
27   </sentenceselection>
28   <quizfeatures version="3">
29     <show>visual</show>
30     <request>gn</request>
31   </quizfeatures>
32 </questiontemplate>
```

The *version* attribute used in several elements identifies what elements may legally appear within other elements. For example, the current version of Bible OL allows the elements `<show>`, `<request>`, `<requestdd>`, and `<dontshow>` within the `<quizfeatures>` element (see line 28 above). If a future

version of Bible OL changes the legal content of `<quizfeatures>`, the `version="3"` string should be changed to `version="4"`.

The top level of the quiz template is the `<questiontemplate>` element. It contains these elements:

Element	Contents
`<desc>`	A *CDATA* string describing the exercise. This string may contain HTML code. This is the text entered under the "Description" tab when creating a quiz template.
`<database>`	The name of the Emdros database (TBD: Or specification file?).
`<properties>`	The name of the Database Localization File (without the *.language*`.prop.json` extension).
`<path>`	This element may occur several times in the file. It describes a component of the passages used for the exercise. This is the data entered under the "Passages" tab when creating a quiz template. In Listing 10.1, lines 6-9 specify Genesis chapter 1 verses 1-2, Genesis chapter 3 (all verses), and the book of Exodus (all chapters).
`<sentenceselection>`	A description of how Bible OL should select sentences. See Section 10.1.
`<quizobjectselection>`	A description of how Bible OL should select sentence units (a.k.a. quiz objects). See Section 10.2.
`<quizfeatures>`	The display features and request features. See Section 10.3.

10.1 `<sentenceselection>`

The `<sentenceselection>` element contains the information entered under the "Sentences" tab when creating a quiz template.

The `<sentenceselection>` element contains these elements:

Element	Contents
`<questionobject>`	The Emdros object type that is used for sentence selection.
`<mql>`	This element is only present if an MQL string is to be used for sentence selection. The element contains the MQL string.
`<featurehandlers>`	This element is not present if an MQL string is to be used for sentence selection. It contains a description of the features used for sentence selection. See Section 10.1.1.
`<useforquizobjects>`	A Boolean value which is *true* if the contents of the `<sentenceselection>` element is also used for sentence unit selection, and is *false* is sentence unit select is specified separately. This is controlled by the check box "Use this for sentence unit selection" under the "Sentences" tab.

10.1.1 <featurehandlers>

The <featurehandlers> element contains descriptions of the Emdros features used for selecting a sentence or a sentence unit. The <featurehandlers> element contains one or more of these elements, each of which describes an Emdros feature and how it is used for selections:

Element	Emdros feature type
<stringfeature>	String. See Section 10.1.1.1.
<integerfeature>	Integer. The selector specifies distinct integer values. See Section 10.1.1.2.
<rangeintegerfeature>	Integer. The selector specifies a range of values for the Emdros feature. See Section 10.1.1.3.
<enumfeature>	Enumeration. See Section 10.1.1.4.
<enumlistfeature>	List of enumeration values. See Section 10.1.1.5.

An implicit logical *AND* is assumed between these selector specifiers, meaning that only objects that match all of the specified selectors are chosen.

10.1.1.1 <stringfeature>

The <stringfeature> element contains a description of how an Emdros feature of type string is used for selecting a sentence or a sentence unit.

The <stringfeature> element contains these elements:

Element	Contents
<name>	The name of the Emdros feature.
<comparator>	The string "equals", "differs", or "matches". If the string is "equals", the Emdros feature must be equal to one of the <value> elements mentioned below; if the string is "differs", the Emdros feature must not be equal to any of the <value> elements mentioned below; if the string is "matches", the Emdros feature must match one of the regular expressions given in the <value> elements below.
<value>	This element may occur several times. It contains a string that is compared to the value of the Emdros feature.

10.1.1.2 <integerfeature>

The <integerfeature> element contains a description of how an Emdros feature of type integer is used for selecting a sentence or a sentence unit.

The <integerfeature> element contains these elements:

Element	Contents
<name>	The name of the Emdros feature.

(Continued...)

Element	Contents
<comparator>	The string "equals" or "differs". If the string is "equals", the Emdros feature must be equal to one of the <value> elements mentioned below; if the string is "differs", the Emdros feature must not be equal to any of the <value> elements mentioned below.
<value>	This element may occur several times. It contains an integer that is compared to the value of the Emdros feature.

10.1.1.3 <rangeintegerfeature>

The <rangeintegerfeature> element contains a description of how an Emdros feature of type integer is used for selecting a sentence or a sentence unit.

The <integerfeature> element contains these elements:

Element	Contents
<name>	The name of the Emdros feature.
<valuelow>	This element is optional. If it is present, it contains an integer. The value of the Emdros feature must be greater than or equal to this value.
<valuehigh>	This element is optional. If it is present, it contains an integer. The value of the Emdros feature must be less than or equal to this value.

10.1.1.4 <enumfeature>

The <enumfeature> element contains a description of how an Emdros feature of enumeration type is used for selecting a sentence or a sentence unit.

The <enumfeature> element contains these elements:

Element	Contents
<name>	The name of the Emdros feature.
<comparator>	The string "equals" or "differs". If the string is "equals", the Emdros feature must be equal to one of the <value> elements mentioned below; if the string is "differs", the Emdros feature must not be equal to any of the <value> elements mentioned below.
<value>	This element may occur several times. It contains string naming an enumeration value that is compared to the value of the Emdros feature.

10.1.1.5 <enumlistfeature>

The <enumlistfeature> element contains a description of how an Emdros feature of type "list of enumeration type" is used for selecting a sentence or a sentence unit.

The <enumlistfeature> element contains these elements:

Element	Contents
<name>	The name of the Emdros feature.
<listvalues>	This element may occur several times. Each specifies a separate selection mechanism. A logical *OR* is assumed between each selection. See below for further information.

The <listvalues> element specifies which enumeration values must occur in the Emdros feature. The <listvalues> element contains these elements:

Element	Contents
yes	This element may occur zero or more times. Each contains an enumeration value that must be present in the Emdros feature.
no	This element may occur zero or more times. Each contains an enumeration value that must not be present in the Emdros feature.

10.2 <quizobjectselection>

The <quizobjectselection> element contains the information entered under the "Sentence Units" tab when creating a quiz template. This element is not present in the quiz template if the <useforquizobjects> element under the <sentenceselection> element is *true*. (See Section 10.1.)

The <quizobjectselection> element may contain the same elements as the <sentenceselection> element except for <useforquizobjects>.

10.3 <quizfeatures>

The <quizfeatures> element lists the display features and request features of the quiz. These must be features of the Emdros object specified in the <questionobject> element of the <quizobjectselection> element (see Section 10.2).

The <quizfeatures> element may contain some or all of the following elements. Each element may occur several times.

Element	Contents
<show>	The name of a display feature.
<request>	The name of a request feature.
<requestdd>	The name of a request feature of string type which should be asked as a multiple-choice question (see Chapter 11).
<dontshow>	The name of a feature that must not be available in the grammar selection box and the grammar information box.

Chapter 11

Multiple-Choice Questions

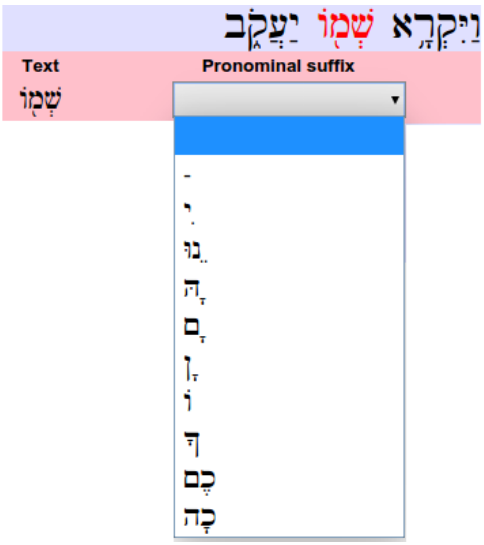
Read this chapter if you need to understand or modify the way Bible OL automatically generates multiple-choice questions for a few features.

If an request feature in an exercise has an enumeration type, the request feature is displayed as a drop-down box and thus becomes a multiple choice question:

But in some cases it is desirable that features of string type are also displayed as multiple choice. For example, in the current version of Bible OL, this is possible for the features *text_nocant_utf8*[1] and *g_prs_utf8*[2].

A separate "Words Database" exists with the information necessary for this to work.

The Words Database is an SQLite3 database containing all possible values for the relevant feature. Bible OL then constructs a drop-down box containing at most ten of the possible values, chosen at random but guaranteed to contain the correct answer:



[1] That is, *Text (no cantillation marks)*.

[2] That is, *Pronominal suffix*.

The strings in the drop-down box are chosen based on three keys from the *objectSettings* of the Database Specification file. (See Section 9.1.1.)

- The name of the Words Database, found under the *alternateshowrequestDb* key.
- An SQL statement that extracts all possible values of the feature. This is found under the *alternateshowrequestSql* key.
- Features to be used as parameters in the SQL statement. These are found under the *additionalfeatures* key.[3]

Let us look at an example from the ETCBC4 database.[4] For the *g_prs_utf8* feature (the pronominal suffix) of the *word* object, the following values are specified under *objectSettings* in the Database Specification File:

Key	Value
additionalfeatures	`["lex"]`
alternateshowrequestDb	`ETCBC4_words.db`
alternateshowrequestSql	`SELECT DISTINCT suffix FROM suffixes,lexsuf,lexemes WHERE lex='%s' AND lexid=lexemes.id AND sufid=suffixes.id`

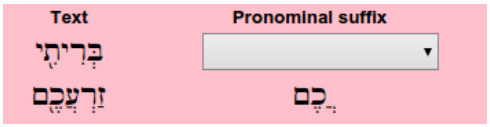
These values will direct Bible OL to replace "%s" in *alternateshowrequestSql* with the value of the *lex* feature retrieved from the Emdros database and then execute the SQL statement on the `ETCBC4_words.db` database.

In the case of the sentence וַיִּקְרָא שְׁמוֹ יַעֲקֹב in the illustration above, the *lex* feature of the word שְׁמוֹ has the value "CM/". Bible OL will therefore execute this SQL statement:

```
SELECT DISTINCT suffix FROM suffixes,lexsuf,lexemes
    WHERE lex='CM/' AND lexid=lexemes.id AND sufid=suffixes.id
```

This will yield a collection of twelve possible pronominal suffixes associated with the *lex* value "CM/". Bible OL will then choose ten of these at random, while still ensuring that the correct answer (וֹ) is among them, and present them in a drop-down box.

If the SQL statement yields only a single value, it is not turned into a drop-down box as with the second word here:



[3] Currently, only one value is allowed in *additionalfeatures*.

[4] For a detailed description of the Words Database for ETCBC4 see Appendix B.

Chapter 12

Data Exchange

Read this chapter if you are going to work with code that accesses the Emdros databases on the server or displays text and exercises in the client.

This chapter describes the data exchange between the client (web browser) and the server.

12.1 Displaying Text

The user requests Bible OL to display a particular passage by accessing a URL in one of these formats:

`http://`*hostname*`/text/show_text/`*dsfname*`/`*book*`/`*chapter*
`http://`*hostname*`/text/show_text/`*dsfname*`/`*book*`/`*chapter*`/`*verse*
`http://`*hostname*`/text/show_text/`*dsfname*`/`*book*`/`*chapter*`/`*firstverse*`/`*lastverse*

The first variant retrieves an entire chapter, the second variant retrieves a single verse, and the third variant retrieves a range of verses.

So, for example, to retrieve Genesis 1 : 2-5 from the ETCBC4-translit database on the server with hostname `bibleol.3bmoodle.dk`, you can use this URL: `http://bibleol.3bmoodle.dk/text/show_text/ETCBC4-translit/Genesis/1/2/5`.

The *dsfname* in the URLs is the first part of the name of the Database Specification File (Section 9.1); so if the *dsfname* is "ETCBC4-translit", the server will access the Database Specification file `ETCBC4-translit.db.json`.

The server will always expand the requested range of verse to contain a complete set of sentences; so a request for Genesis 1 : 16-17 will automatically be expanded to include verse 18, because verses 17 and 18 comprise a single sentence.

When the server has interpreted the components of the URL, it queries the relevant Emdros database, and based on the result it generates an HTML document containing little more than the relevant headers, HTML code to lay out the menu, and these JavaScript variables:

Variable	Contents
useToolTip	A Boolean value of *true* if the grammar information box should be displayed as a tooltip under the mouse, *false* if the grammar information box should be displayed at the right side of the browser window. This value is configurable on a per-user basis, but currently there is no user interface to change its value.
configuration	A JavaScript object whose value is the contents of the Database Specification File (Section 9.1).

(Continued...)

Variable	Contents
localization	A JavaScript object whose value is the contents of the Database Localization File (Section 9.3).
typeinfo	A JavaScript object whose value is the contents of the Database Type Information File (Section 9.4).
site_url	The base part of the URL of the website. (For example, `http://bibleol.3bmoodle.dk/`.)
dictionaries	The text to display, including grammar information. This is a JavaScript object in a format defined in Chapter 13.
quizdata	This variable is *null*, indicating that we are displaying text, not running an exercise.

Included in the HTML document which the server sends to the client is a link to a number of CSS and JavaScript files, including `ol.js` which contains the main piece of code that is to run on the client.

When the client (the web browser) has read the HTML file the JavaScript code directs it to construct the visual appearance of the text. This involves building the central text layout, adding grammatical information to each word, phrase, clause, etc., and constructing the grammar selection box and the grammar information box.

12.2 Running an Exercise

The user requests Bible OL to start a particular exercise by accessing a URL in one of these formats:

`http://`*hostname*`/text/show_quiz?quiz=`*excercisename*`&count=`*numberOfQuestions*
`http://`*hostname*`/text/show_quiz_univ?quiz=`*excercisename*`&count=`*numberOfQuestions*

The first variant executes an exercise based on a pre-defined set of Bible passages (which are stored as part of the exercise file); the second variant asks the user which Bible passages to use, and then starts the exercise. In both cases the *exercisename* is the path name (relative to the `quizzes` directory) of the file containing the exercise; the *numberOfQuestions* is a positive integer indicating the maximum number of questions to ask in the exercise.[1]

So, for example, to run an exercise consisting of ten questions from the exercise file `Nestle 1904/demo/case.3et` on the server with hostname `bibleol.3bmoodle.dk`, you can use this URL: `http://bibleol.3bmoodle.dk/text/show_quiz?quiz=Nestle%201904/demo/case.3et&count=10`. This will use the pre-defined set of Bible passages stored with the exercise file.

If the user chooses the second variant of URL, the server will display a tree of books of the Bible. The user can open book nodes to display a list chapters, and they can open chapter nodes to display a list of verses. The client retrieves the number of chapters in each book and the number verses in each chapter as needed using AJAX requests from the *jstree* package (see page 21). Once the user is click the "Start quiz" button, the client sends an HTTP POST request to the server containing the exercise file name, the number of questions to ask, and the list of selected passages.

The server uses information in the exercise file and the list of passages (either the list specified by the user or the pre-defined list for the exercise) to generate a query for the relevant Emdros database. Based on the result of the query, the server generates an HTML document containing little more than the relevant headers, HTML code to lay out the menu, and these JavaScript variables:

[1] If `count` is omitted or if the number is illegal, five questions will be asked.

Variable	Contents
useToolTip	The same as in Section 12.1.
configuration	The same as in Section 12.1.
localization	The same as in Section 12.1.
typeinfo	The same as in Section 12.1.
site_url	The same as in Section 12.1.
dictionaries	For each question, this variable contains the text to display, including grammar information. This is a JavaScript object in a format defined in Chapter 13.
quizdata	A JavaScript object containing information about the exercise being run. This includes information about the display features and the request features and their (correct) values. This variable is described in Chapter 14.

Included in the HTML document which the server sends to the client is a link to a number of CSS and JavaScript files, including `ol.js` which contains the main piece of code that is to run on the client.

When the client (the web browser) has read the HTML file the JavaScript code directs it to construct the visual appearance of the exercises. This involves building the central text layout, adding grammatical information to each word, phrase, clause, etc., building the question/answer panel, and constructing the grammar selection box and the grammar information box.

As the user answers the questions, the client keeps track of the answers, but no communication with the server takes place before the user presses the "Finish" button. What happens when the "Finish" button is pressed, depends on whether the user is logged in or not. If the user is logged in, the client sends the result to the server URL `http://`*hostname*`/statistics/update_stat`, which updates the statistics for the user; if the user not logged in, the client instructs the server to display the *Select a Quiz* web page.

Chapter 13

The *dictionaries* Variable

Read this chapter if you are going to work with code that accesses the Emdros databases on the server or displays text and exercises in the client.

All information about the text to display, the associated feature values, and the phrase, clause, and sentence structure of the text is communicated between the server and the client in the JavaScript variable *dictionaries*.

This means that the primary task of the server is to convert data from the Emdros database into a format that can be stored in the *dictionaries* variable, and the primary task of the client is to convert the contents of the *dictionaries* variable into displayed text.

In the PHP files executed by the server, the value is described by the *Dictionary* class in the file `myapp/libraries/Dictionary.php`. In the TypeScript files executed by the client, the value is described by *DictionaryIf* interface in the file `ts/dictionary.ts`.

Note: There is also a class in the client called *Dictionary*. This is not the same as the *Dictionary* class in the server (although the two are related, as we shall see in Section 19.2).

13.1 The TypeScript *DictionaryIf* and PHP *Dictionary* Classes

There is a one-to-one mapping of classes and data fields in the server and in the client, but the names differ a little, as you can see from the following.

This is a simplified overview of the relevant TypeScript definitions in the client:

```
interface DictionaryIf {
    sentenceSets : MonadSet[];
    monadObjects: MonadObject[][][];
    bookTitle : string;
}

interface MonadSet {
    segments : MonadPair[];
}

interface MonadPair {
    low : number;
    high : number;
}

interface MonadObject {
    mo : MatchedObject;
    children_idds : number[];
}

interface MatchedObject {
```

```
    id_d : number;
    name : string;
    monadset : MonadSet;
    features : {[key : string] : string;};
    sheaf : any;
}

interface SingleMonadObject extends MonadObject {
    text : string;
    suffix : string;
    bcv : string[];
    sameAsNext : boolean[];
    sameAsPrev : boolean[];
    pics : number[];
    urls : any[];
}

interface MultipleMonadObject extends MonadObject {
}
```

This is a simplified overview of the corresponding PHP definitions in the server:

```
class Dictionary {
    public $sentenceSets; // An array of OlMonadset objects
    public $monadObjects; // An array of arrays of arrays of MonadObject objects
    public $bookTitle;
}

class OlMonadSet implements Iterator {
    public $segments;   // An array of MonadPair objects
}


class MonadPair {
    public $low;
    public $high;
}

abstract class MonadObject {
    public $mo;         // An OlMatchedObject object
    public $children_idds;
}

class OlMatchedObject {
    public $id_d;
    public $name;
    public $monadset;   // An OlMonadSet object
    public $features;   // Maps feature name to feature value
    public $sheaf;
}


class SingleMonadObject extends MonadObject {
    public $text;
    public $suffix;
    public $bcv;        // An array of strings or integers
    public $sameAsNext; // An array of Booleans
    public $sameAsPrev; // An array of Booleans
    public $pics;       // An array of integers
    public $urls;       // An array of string pairs
}

class MultipleMonadObject extends MonadObject {
```

```
}
```

In the following text, the names from the client definitions are used.
The fields of the *DictionaryIf* interface are:

Field	Contents
sentenceSets	Each element in this array specifies a lump of text, defined by the constituent monads. When using the "Display text" feature of Bible OL, *sentenceSets* contains only a single lump of text, and the array therefore has a single element. When running an exercise, each question uses a separate lump of text, and there are as many entries in the array as there are questions in the exercise. If, for example *sentenceSets[2]* has the value `{"segments": [{"low": 8868, "high": 8877}]},` it means that the third question (index 2) is based on monads 8868-8877 from the Emdros database.
monadObjects	An array of arrays of arrays of *MonadObject* objects. The first index of this array corresponds to the index in the *sentenceSets* array. The second index selects the level in the grammatical hierarchy (for example, word – phrase – clause – sentence). At the third level we find the actual objects. For example, *monadObjects[0][2][4]* is the *MonadObject* that describes the fifth (index 4) clause (index 2) of the first (index 0) lump of text.
bookTitle	The title of the current book of the Bible. It is used by the client to display an appropriate heading on the webpage when displaying text.

The grammatical information about each object is structured in a grammatical hierarchy. At the lowest level we have words. The names of the objects above the words are named phrase, clause, and sentence in ETCBC4; they are named clause level 2, clause level 1, and sentence in nestle1904. The names and the depth of the hierarchy can be chosen freely by the database, although the user interface can only display a limited number of levels.

Above the top level, all sentences in a lump of text are grouped together in a so-called "patriarch" object.

Each Emdros object is described by a *MonadObject* int the *monadObjects* array mentioned above. The second index of *monadObjects* identifies the level in the grammatical hierarchy. *MonadObject* is an abstract class, its concrete subclasses are *SingleMonadObject* and *MultipleMonadObject*.

At the lowest level (the word level) the *MonadObjects* belong to the class *SingleMonadObject,* which represents Emdros objects that correspond to a single monad. At the higher levels in the grammatical hierarchy, the *MonadObjects* belong to the class *MultipleMonadObject,* which represents Emdros objects that correspond to a multiple monads.

Consider, for example, the middle sentence of Genesis 1 : 7:[1]

wayyavdēl bên hammayim ʔašer mittaḥat lārāqîₐʕ ûvên hammayim ʔašer mēʕal lārāqîₐʕ[2]

[1] I am using a transliterated text here because a left-to-right orientation will make the following illustrations easier to read.

[2] In English: "and separated the waters that were under the expanse from the waters that were above the expanse."

At grammatical level 0, the words are assigned these monads:

wa-	yyavdēl	bên	ha-	mmayim	ʔᵃšer	mi-	ttaḥat	lā-	-	rāqîₐʕ
101	102	103	104	105	106	107	108	109	110	111

û-	vên	ha-	mmayim	ʔᵃšer	mē-	ʕal	lā-	-	rāqîₐʕ
112	113	114	115	116	117	118	119	120	121

Assuming that the *dictionaries* variable contains just this one lump of text (that is, *sentenceSet* contains only one element), the 21 words are represented as *SingleMonadObjects* in *monadObjects[0][0][0..20]* in the *dictionaries* variable. (The "words" with monads 110 and 120 a null forms of the definite article.)

At grammatical level 1, the words are grouped into phrases thus:

wa-	yyavdēl	bên hammayim ûvên hammayim	ʔᵃšer	mittaḥat lārāqîₐʕ	ʔᵃšer	mēʕal lārāqîₐʕ
101	102	103-105, 112-115	106	107-111	116	117-121

Note how the third phrase consists of two lumps of contiguous monads. The seven phrases are represented as *MultipleMonadObjects* in *monadObjects[0][1][0..6]* in the *dictionaries* variable.

At grammatical level 2, the phrases are grouped into clauses thus:

wayyavdēl bên hammayim ûvên hammayim	ʔᵃšer mittaḥat lārāqîₐʕ	ʔᵃšer mēʕal lārāqîₐʕ
101-105, 112-115	106-111	116-121

The three clauses are represented as *MultipleMonadObjects* in *monadObjects[0][2][0..2]* in the *dictionaries* variable.

At grammatical level 3, the clauses are grouped into sentences thus:

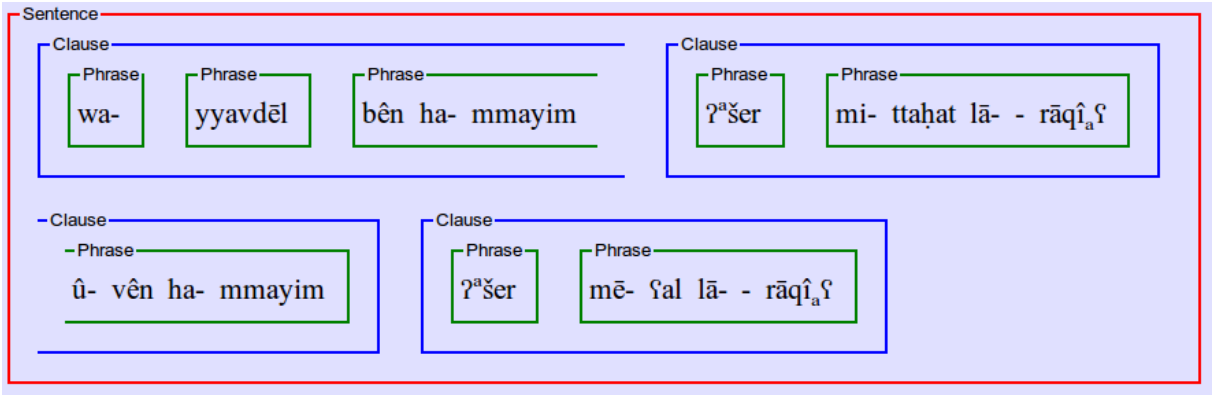
wayyavdēl bên hammayim ʔᵃšer mittaḥat lārāqîₐʕ ûvên hammayim ʔᵃšer mēʕal lārāqîₐʕ

101-121

In this example, only a single sentence is present. This sentence is represented as a *MultipleMonadObject* in *monadObjects[0][3][0]* in the *dictionaries* variable.

Finally, at grammatical level 4, the sentences are grouped into a single patriarch object. In this example the patriarch contains the same monads as the sentence object. The patriarch is represented as a *MultipleMonadObject* in *monadObjects[0][4][0]* in the *dictionaries* variable.

In Bible OL the grammatical hierarchy can be displayed thus:



Note how the frames around the split phrase and clause are drawn.

13.2 The *MonadObject* Class and Its Subclasses

The previous section describes how objects of class *MonadObject* are used to represent Emdros objects in the *monadObjects* field in the *dictionaries* variable. (Strictly speaking, *MonadObject* is an interface in TypeScript and an abstract class in PHP, but that is irrelevant to the following discussion.)

The *MonadObject* class has these members:

Field	Contents
mo	An object of class *MatchedObject* (*OlMatchedObject* in PHP). A *MatchedObject* is a representation of data about a single Emdros object. Details are given in Section 13.3.
children_idds	An array of integers containing the ID_Ds of the constituent Emdros objects at a lower level in the grammar hierarchy. Consider, for example, the three clauses at grammatical level 2 in the example on page 61. The clause "ʔᵃšer mittaḥat lārāqîₐʕ" contains the two phrases "ʔᵃšer" and "mittaḥat lārāqîₐʕ". If these two phrases have ID_Ds 357 and 362, then *children_idds* of the clause will be the array [357, 362].

At the lowest level (the word level) in the grammar hierarchy, where every Emdros object corresponds to a single monad, the *SingleMonadObject* subclass of *MonadObject* is used to represent an Emdros object.

In addition to the fields mentioned above, a *SingleMonadObject* has these members:

Field	Contents
text	A string containing the text used to display the word.
suffix	A string containing the continuation feature, if any, of the word. (See Section 8.2.2.)
bcv	An array containing the Bible reference for the verse containing the word. For a word in Genesis 1 : 7, *bcv* is the array ["Genesis", 1, 7].
sameAsNext	An array of Booleans. *sameAsNext[0]* is *true* if this word belongs to the same book as the next word; *sameAsNext[1]* is *true* if this word belongs to the same chapter as the next word; *sameAsNext[2]* is *true* if this word belongs to the same verse as the next word. For the last word in a collection, all three Booleans are *false*.
sameAsPrev	An array of Booleans. *sameAsPrev[0]* is *true* if this word belongs to the same book as the previous word; *sameAsPrev[1]* is *true* if this word belongs to the same chapter as the previous word; *sameAsPrev[2]* is *true* if this word belongs to the same verse as the previous word. For the first word in a collection, all three Booleans are *false*.
pics	An array of integers identifying pictures on the resource website (see Section 20.2) that are relevant for this word. The first three values in the array are the book number, chapter number, and verse number. The remaining elements are IDs of pictures on the resource website.
urls	An array of references to URLs that are relevant for the word. Each entry in the array is itself an array with two element: the URL and a string identifying the icon to show in Bible OL. If, for example an element in *url* has the value ["http://example.com/here.html", "v"], Bible OL will show a hyperlink to `http://example.com/here.html` in the form of a V icon. For more information see Section 20.2.

A *MultipleMonadObject* has no fields other than those defined for *MonadObject*.

13.3 The *MatchedObject* Class

A *MatchedObject* (called an *OlMatchedObject* in PHP) is a representation of data about a single Emdros object. It has the following members:

Field	Contents
id_d	The ID_D of the Emdros object
name	The name of the Emdros object type.

(Continued...)

Field	Contents

monadset	A *MonadSet* (*OlMonadSet* in PHP) object listing the monads belonging to this object. For example, the first clause at grammatical level 2 in the example on page 61 is represented by this value:
	`"monadset": {`
	`"segments": [`
	`{`
	`"low": 101,`
	`"high": 105`
	`},`
	`{`
	`"low": 112,`
	`"high": 115`
	`}`
	`]`
	`}`

features	An associative array mapping feature name to feature value.
sheaf	Always null. (This is used by complex MQL searches that should not occur in Bible OL.)

Chapter 14

The *quizdata* Variable

Read this chapter if you are going to work with code that generates exercises on the server or displays exercises in the client.

All information about the questions and correct answers to an exercise is communicated between the server and the client in the JavaScript variable *quizdata*.

In the server, the data is generated by the class *Quiz_data* in the file `myapp/libraries/Quiz_data.php`; in the client the data is described in the TypeScript *QuizData* interface in the file `ts/quizdata.ts`. Th

14.1 The TypeScript *QuizData* and PHP *Quiz_data* Classes

There is a one-to-one mapping of classes and data fields in the server and in the client, but the names differ a little, as you can see from the following.

This is a simplified overview of the relevant TypeScript definitions in the client:

```
interface QuizData {
    quizid : number;
    quizFeatures : ExtendedQuizFeatures;
    desc : string;
    monad2Id : number[];
    id2FeatVal : string[][];
}

interface ExtendedQuizFeatures {
    showFeatures : string[];
    requestFeatures : {name : string; usedropdown : boolean; }[];
    dontShowFeatures : string[];
    objectType : string;
    dontShow : boolean;
    useVirtualKeyboard : boolean;
}
```

This is a (very) simplified overview of the corresponding PHP definitions in the server:

```
class Quiz_data {
    public $quizid;
    public $quizFeatures;     // An ExtendedQuizFeatures object
    public $desc;
    public $monad2Id;         // An array of integers
    public $id2FeatVal;       // An array of arrays of strings
}

class ExtendedQuizFeatures {
```

```
    public $showFeatures;     // An array of strings
    public $requestFeatures;  // An array of string/Boolean pairs
    public $dontShowFeatures; // An array of strings
    public $objectType;
    public $dontShow;
    public $useVirtualKeyboard;
}
```

In the following text, the names from the client definitions are used.
The fields of the *QuizData* interface are:

Field	Contents
quizid	An integer used to identify the entry in the *sta_quiz* table in the user database (see Section 17.11.1) where statistics about this exercise is to be stored. If the user is not logged in, *quizid* is -1 and statistics will not be stored.
quizFeatures	An object of class *ExtendedQuizFeatures*. It contains information about how the exercise should be presented to the user. Details are given below.
desc	The description of the exercise from the quiz template file.
monad2Id	An array that maps monads to the ID_Ds of quiz object. If a quiz object covers more than one word, several entries in this array will have the same value. (Note that this assumes that a word is not part of more than one quiz object.)
id2FeatVal	An array of arrays of strings. For each quiz object, this array holds information about the values of the display features and the correct values of the request features. If, for example, a quiz object has an ID_D of 1234, and one of the display or request features is *case* with the value *genitive*, then *id2FeatVal[1234]['case']* will have the value "genitive".

The *ExtendedQuizFeatures* interface has these members:

Field	Contents
showFeatures	An array of strings containing the names of the display features.

requestFeatures	An array of objects describing the request features. Each object has this layout:
	`{`
	`  name : string;`
	`  usedropdown : boolean;`
	`}`
	The *name* field contains the name of the request feature, the *usedropdown* field is *true* if the request feature is of type string and the question should be asked as a multiple choice question (see Section 11).

dontShowFeatures	Array of strings naming features that must not be available in the grammar selection box and the grammar information box.
objectType	The Emdros type of the quiz objects.
dontShow	*True* if the quiz objects should be replaced with (1), (2), (3), etc. in the displayed text.
useVirtualKeyboard	*True* if the client should display a virtual keyboard to facilitate the typing of text in a foreign alphabet.

The *quizFeatures* field of the *QuizData* interface has the additional fields *useDropdown, additionalFeatures,* and *allFeatures.* They are not used by the client software.

Chapter 15

The CodeIgniter Framework

Read this chapter if you are going to understand or modify the server code.

Bible OL uses a PHP framework known as *CodeIgniter*, which provides a simple decoding of URLs, forces a model-view-controller approach to the software structure, and provides a large library for performing a number of tasks.

If you are going to modify server code, you will need a good understanding of how CodeIgniter works, and you should therefore read the documentation at http://www.codeigniter.com/docs. The following sections provide a few examples of what CodeIgniter can do for the programmer.

15.1 URL Decoding

When a user accesses a URL such as, for example, `http://website/aaaaa/bbb`, this will cause CodeIgniter to call the PHP function *bbb* in the class *Ctrl_aaaaa*, which is located in the file *ctrl_aaaaa.php*.

It is also possible to access these functions from the shell command line on the server. The shell command

```
php index.php aaaaa bbb
```

is equivalent to accessing `http://website/aaaaa/bbb`.

If the function name is omitted, it defaults to *index*.

15.2 Model-view-controller Structure

It is customary in many large programming projects to split functionality into three groups:

- Models, which are responsible for providing the data that is to be displayed to the user.
- Views, which handle the actual layout on the computer screen.
- Controllers, which handle the flow of data between the models and the views.

CodeIgniter makes structuring PHP code into model-view-controller groups easy.

Continuing with the example from the previous section, the function *bbb* contains the *controller* code.

This controller function may load one or model *models*. If, for example, *bbb* executes this code:

```
$this->load->model('mod_users');
$this->mod_users->get_user_by_id(8);
```

the model class *Mod_users* is loaded from the file *mod_users.php* and the function *get_user_by_id(8)* is called in that class. The *Mod_users* class handles the data exchange with the underlying user database.

Once the controller has retrieved the relevant data, it may load a *view* class which handles the generation of the HTML code presented to the browser. If, for example, *bbb* executes this code:

```
$this->load->view('view_main_page', array('name' => $user_name,
                                          'email' => $user_email));
```

the view file *view_main_page.php* will be loaded and the variables `$user_name` and `$user_email` will be transferred to the file. The code in the view file (mostly HTML) will then be sent to the browser.

15.3 Library Functions

The CodeIgniter library provides a large set of library functions. One of the most important is a set of functions that enable easy and safe construction of SQL statements. For example, instead of the PHP/SQL statement

```
SELECT * from {$db_prefix}user WHERE name=$username AND age=$userage ORDER BY id;
```

you can write this PHP code:

```
$this->db->select('*')
     ->from('user')
     ->where('name', $username)
     ->where('age', $userage)
     ->order_by('id');
```

15.4 Adding Code

Almost all PHP code resides in the directory `myapp`. Controller, model, and view classes are found in the directories `myapp/controllers`, `myapp/models`, and `myapp/views`, respectively. Functions can be added to existing classes, or files containing new classes can be added to the subdirectories.

The code for CodeIgniter itself resides in the directory `CodeIgniter`.

CodeIgniter is configured through definitions in the files located in the directory `myapp/config`. I have added a file `ol.php` to this directory containing these configuration variables:

Index in $config	Value
pw_salt	Salt for storing the password in the user database.
mql_driver	Set to `'native'` to select a built-in MQL driver. Set to or `'extern'` to run MQL commands external to the PHP interpreter.
mail_sender_address	The sender address on email sent by Bible OL to registered users.
mail_sender_name	The sender name on email sent by Bible OL to registered users.
google_client_id	Used by Google login (see Section 16.4).
google_client_secret	Used by Google login (see Section 16.4).

Chapter 16

Server Code

Read this chapter if you are going to understand or modify the server code.

This chapter describes some of the techniques and tools you will find in the server code. The server code is written in PHP, and almost the server code resides in the directory `myapp`. The server code uses the CodeIgniter framework (see Section 15), and a good understanding of how CodeIgniter works is essential to understanding the server code.

16.1 Models, Views, and Controllers

Almost all PHP code used by the server resides in the directory `myapp`. The directories `models`, `views`, and `controllers` hold the main components of the MVC structure supported by CodeIgniter.

The following controllers exist:

Name	Function
classes	Manages classes (that is, groups of users).
config	Allows users to change their font preferences.
file_manager	Management of files and directories in the `quizzes` directory.
google	Handles login using a Google account.
login	Handles login using the local user database.
main_page	Displays the main page.
maketypeinfo	This controller can only be accessed from the command line, not through the web interface. It is used to create the Database Type Information File as described in Section 9.4.
pic2db	This controller can only be accessed from the command line, not through the web interface. It is used to retrieve information from the resource website (see Section 20.2).
privacy	Displays the privacy policy.
statistics	Updates and displays statistics about the exercises executed by the user.
text	Displays text or exercises. Also handles editing of exercises.
transfer_stat	This controller can only be accessed from the command line, not through the web interface. It was used to transfer statistics to the *Learning Journey* website. This controller is now obsolete, because *Learning Journey* accesses the statistical data directly from the Bible OL database.

(Continued...)

Name	Function
userclass	Manages a user's relationship to a class.
users	Manages users.

The following models exist:

Name	Function
mod_askemdros	Retrieves Emdros-related data.
mod_classdir	Manages class permissions for an exercise directory.
mod_classes	Manages classes (that is, groups of users).
mod_config	Manages users' font preferences.
mod_intro_text	Generates the text on the front page.
mod_quizpath	Exercise directory operations.
mod_statistics	Manages user statistics.
mod_userclass	Manages a user's relationship to a class.
mod_users	Manages users.

There is no reason to go through the various views here. Their function is best learned by looking for calls like `$this->load->view(...)` in the controller code.

In addition to the model/view/controller classes, the following modules are worth noting:[1]

File (add "`.php`" to name)	Contents
`core/MY_Controller`	Customized version of CodeIgniter's *CI_Controller* class.
`helpers/quiztemplate_helper`	XML parser for exercise files.
`helpers/sheaf_helper`	Classes that model data in Emdros replies.
`helpers/sheaf_xml_helper`	XML parser for Emdros replies.
`helpers/xmlhandler_helper`	Superclass and functions for XML parser.
`libraries/DB_config`	Classes for handling Emdros database description files.
`libraries/Dictionary`	The *Dictionary* class (see Chapter 13).
`libraries/include/dataexception.inc`	Exception classes.
`libraries/include/monadobject.inc`	The *MonadObject* class and its subclasses (see Section 13.2).
`libraries/include/typeinfo.inc`	The *TypeInfo* class (see Section 9.4).
`libraries/Mql/Mql`	The *Mql* class which handles MQL requests (see Section 16.3).
`libraries/Mql/drivers/Mql_extern`	Driver for executing MQL commands through an external MQL command.

(Continued...)

[1] Note that the list is not complete. Consult the comments in the individual files for more information.

File (add ".php" to name)	Contents
libraries/Mql/drivers/Mql_native	Driver for executing MQL commands through an MQL library in PHP.
libraries/picdb	Class for retrieving picture references and URL references from the resource website (see Section 20.2).
libraries/Quiz_data	The *Quiz_data* class and associated functions and class (see Chapter 14).
libraries/Suggest_answers	Class for accessing the Words Database (see Chapter 11).
libraries/Universe_tree	Classes used together with *jstree* (see page 21) to display a hierarchy of books, chapters, and verses of the Bible.

16.2 PHP Parameter Type Hinting

Unlike better programming languages, PHP does not provide any type safety. In a declaration such as this:

```
function foo($x) { ... }
```

there is no way to tell if *$x* should be an integer, a string, a Boolean, an array, or perhaps an instance of some class.

Recent versions of PHP have tried to remedy this by introducing *type hinting*[2]. Using type hinting, you can specify that an argument should be, for example, an array or an instance of a class:

```
function foo(Xyz $x) { ... } // $x must be an object of class Xyz
function bar(array $x) { ... } // $x must be an array
```

However, this does not work for simple types such as integers or strings.

The Bible OL code uses a special technique which involves catching error messages and interpreting them to achieve better (but not perfect) type safety. In Bible OL server code you can specify simple types as type hints:

```
function foo(integer $x) { ... }
function bar(string $x) { ... }
function pip(boolean $x) { ... }
```

If you need to be able to handle two different types you can do this:

```
function foo(integer__OR__string $x) { ... } // $x must be an integer or a string
function bar(array__OR__null $x) { ... } // $x must be an array or null
```

These type hints are handled automatically in the file `typehint.inc.php`.

16.3 MQL Requests in the Server

The server code can be configured to execute MQL requests in one of two ways:

- Adding an MQL library to PHP and calling the MQL API directly from PHP.
- Executing the command line version of MQL from within PHP code.

[2] See http://php.net/manual/en/language.oop5.typehinting.php.

A driver layer in Bible OL protects the programmer for having to worry about this in most cases; this is described in Section 16.3.1. Information about the two drivers are found in Sections 16.3.2 and 16.3.3.

16.3.1 The MQL Interface to Bible OL

The Driver Library mechanism of CodeIgniter is used to hide the MQL implementation from the programmer in most situations. The programmer must specify the desired way to interact with MQL by setting *$config['mql_driver']* in `myapp/config/ol.php` to either *'native'* (for using the MQL PHP library) or *'extern '* (for using an external MQL command).

The programmer can access MQL in the following way. First, the appropriate MQL driver must be loaded:

```
$this->load->driver('mql',array('db' => $this->db_config->emdros_db,
                                'driver' => $this->config->item('mql_driver')));
```

This is typically done in the *setup* function of the *mod_askemdros* module. Here, *$this->db_config->emdros_db* is the name of the Emdros database.

After this, MQL requests can be executed like in this example:

```
$emdros_data = $this->mql->exec("SELECT ALL OBJECTS WHERE [word sp=subs GET text] GOqxqxqx");
```

It is important that each Emdros command be terminated by "GOqxqxqx" rather than simly "GO". The reason is that the MQL driver needs to split a string of several Emdros commands into individual commands. It does this by looking for the string "GOqxqxqx". If the string "GO" had been used instead, the occurrence of a "GO" inside an MQL command would cause the command splitting to fail. The string "GOqxqxqx" is chosen because it is highly unlikely to occur inside an MQL command.

The call to the *exec* function executes the MQL command and returns the result as an array of *TableOrSheaf* objects. The *TableOrSheaf* class is defined in `myapp/helpers/sheaf_helper.php`.

If the result of the MQL query is a table, the *get_table* function of *TableOrSheaf* will return the table as an *OlTable* object. This object has functions such as *rows, cols, get_header* and *get_cell* which allow you to access various parts of the table.

If the result of the MQL query is a sheaf or a flat sheaf, the *get_sheaf* function of *TableOrSheaf* will return the table as an *OlSheaf* object. This object has functions such as *get_straws, get_first_straw,* and *number_of_straws* which allow you to access the straws within the sheaf.

If you know that the MQL request will return a sheaf and you are only interested in the monads of that sheaf, the so-called "quick harvest" method can be used. In this case you must add a Boolean argument with the value *true* to the call to *exec*:

```
$emdros_data = $this->mql->exec("SELECT ALL OBJECTS WHERE [word sp=subs] GOqxqxqx", true);
```

As before, *exec* returns an array of *TableOrSheaf* objects, but in this case all the objects represent sheafs. As before, calling *get_sheaf* on one of these objects returns an *OlSheaf* object, but in this case the *OlSheaf* only contains *OlMonadSet* objects, and the functions for accessing straws do not work. The function *has_monadset* returns *true* if any *OlMonadSets* are available; and the function *get_monadset* returns an array of *OlMonadSets*.

16.3.2 Driver for Native MQL

The driver for native MQL assumes that MQL support has been added to PHP. Section 6.1.1 explains how to do this.

Using the native MQL driver, the MQL API is available directly from within PHP. A C++ version of the MQL API is described in the Emdros Programmer's Reference Guide[3].

16.3.3 Driver for External MQL

The driver for external MQL relies on the existence of an MQL command line tool on the server. The driver (located in file `myapp/libraries/Mql/drivers/Mql_extern.php`) contains this variable definition

```
private $command_line = '/usr/local/bin/mql --xml';
```

If the MQL command line program is located in some other directory, this variable definition must be changed accordingly.

Note that there is almost no error reporting from MQL when the external MQL command is used.

16.4 Google Login

The server provides two different login mechanism. One uses a local list of users (see Section 17.1), the other relies on a user's Google login. Google's description of how login works can be found here: https://developers.google.com/accounts/docs/OAuth2WebServer. The following is a brief description of the mechanism as it is set up on the Bible OL installation that runs at http://bibleol.3bmoodle.dk.

Step 1. The Browser Sends Authentication Request to Google

When a user clicks "Sign in with Google+" on the login page, the browser sends an HTTP GET request to `https://accounts.google.com/o/oauth2/auth` with the following GET parameters:

Name	Value
response_type	`'code'`
client_id	Our Google client ID, configured in the file `myapp/config/ol.php`.
redirect_uri	`'http://bibleol.3bmoodle.dk/google/callback'`
scope	`'https://www.googleapis.com/auth/userinfo.profile https://www.googleapis.com/auth/userinfo.email'`
state	A random value, stored in CodeIgniter's session mechanism.

Step 2. Google Responds

Google checks if the user can be logged in to Bible OL. Google then sends a response to the browser, directing it to send a new HTTP GET command to the *redirect_uri* specified in the request above. The parameters for the new GET request depend on wheter Google approved access or not.

Step 3. Browser Sends Authentication Information to the Bible OL Server

As directed by the response from Google, the browser sends a HTTP GET request to `http://bibleol.3bmoodle.dk/google/callback` with the following GET parameters:

[3] http://emdros.org/progref/current

Name	Value
error	Error information if acceess is denied. If access is approved, this parameter is not present.
state	The value of the *state* parameter from Step 1.
code	An authentication code generated by Google.

Step 4. The Bible OL Server Requests an Access Token from Google

In Bible OL, the request from the browser is handled by the function *callback* in the *Ctrl_google* controller (in the file `myapp/controllers/ctrl_google.php`).

If Google approved access, Bible OL does not respond immediately to the client, but sends an HTTP POST request to `https://accounts.google.com/o/oauth2/token` with the following POST parameters:

Name	Value
code	The authentication code received in the *code* parameter in Step 3.
client_id	Our Google client ID, configured in the file `myapp/config/ol.php`.
client_secret	Our Google client secret, configured in the file `myapp/config/ol.php`.
redirect_uri	`'http://bibleol.3bmoodle.dk/google/callback'`
grant_type	`'authorization_code'`

In response to the HTTP POST request, Google replies with a JSON string containing an *access_token*.

Step 5. The Bible OL Server Requests User Information from Google

The Bible OL server sends an HTTP GET request to `https://www.googleapis.com/oauth2/v1/userinfo` with the following GET parameter:

Name	Value
access_token	The *access_token* received from Google in Step 4.

In response to this request, Google replies with a JSON string containing *id, given_name, family_name,* and *email* for the user. If this is the first time the user logs in to Bible OL, the user information is stored in the user table (Section 17.1) of the user database. Bible OL generates a username by concatenating the string "ggl_" with the user *id* received from Google.

Step 6. User is Logged In

The Bible OL server now responds to the request sent by the browser in Step 3. The response consists of web page informing the user that they are now logged in to the system.

Chapter 17

User Database

Read this chapter if you are going to set up a server installation or if you are going to use the user database

The user database is a MySQL database the contains information about the users and classes registered on the system.

The database contains these tables:

alphabet	font	sta_requestfeature
bible_refs	personal_font	sta_universe
bible_urls	sta_displayfeature	user
class	sta_question	userclass
classexercise	sta_quiz	userconfig
exercisedir	sta_quiztemplate	

The names are typically prefixed by a common text string found in *$db['default']['dbprefix']* in the file `myapp/config/database.php`. If that value is *'ol_'*, the database tables will be named *ol_alphabet, ol_bible_refs* etc.

The tables are described in the following sections.

17.1 The *user* Table

The *user* table contains information of each registered user. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying the user.
first_name	Text	User's first name.
last_name	Text	User's last name
username	Text	Name used when logging in.
password	Text	Encrypted password.
reset	Text	Password reset code.
reset_time	Integer	UNIX time when password reset code was issued.

(Continued...)

Column	Type	Contents
google_login	Boolean	User uses Google login.
isadmin	Boolean	User is an administrator.
may_see_wivu	Boolean	User is allowed to see the entire WIVU database.
email	Text	User's email address.

Local users have their username and password stored in this table. Their *google_login* field is set to *false* (0).

Google users have usernames such as "ggl_106440263559736360192", where 106440263559736-360192 is the user ID provided by Google. Their *google_login* field is set to *true* (1).

The *may_see_wivu* was used previously for the WIVU database, which was available in two versions: the entire Old Testament and a subset containing about 20% of the text.

17.2 The *class* Table

The *class* table contains information of each class (as in "school class", a group of students). It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying the class.
classname	Text	Name of the class.
password	Text	An optional password required when a student enrols in a class.
enrol_before	Date	An optional deadline for enrolment.

17.3 The *userclass* Table

An entry in the *userclass* table indicates that a particular user is a member of a particular class. The table has these fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
userid	Integer	The *id* field from an entry in the *user* table.
classid	Integer	The *id* field from an entry in the *class* table.

17.4 The *userconfig* Table

The *userconfig* table holds information about the configuration for a particular user. Currently, only one option is available, and there is no user interface for configuring it. The table has these fields:

Column	Type	Contents
user_id	Integer	The *id* field from an entry in the *user* table.
usetooltip	Boolean	*True* if the user wants the grammar information box to work as a tooltip instead of having a fixed position on the display.

17.5 The *alphabet* Table

The *alphabet* table contains a list of the foreign alphabets used by Bible OL. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying the alphabet.
name	Text	The internal name of the alphabet.
direction	Text	"rtl" for right-to-left text, "ltr" for left-to-right text.
sample	Text	A sample text in the alphabet. This will be displayed when the user chooses fonts.
english	Text	The English name of the alphabet.

17.6 The *font* Table

The *font* table contains the users' font preferences for various alphabets. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
user_id	Integer	The *id* field from an entry in the *user* table.
alphabet_id	Integer	The *id* field from an entry in the *alphabet* table.
font_family	Text	A comma-separated string of font names.
text_size	Integer	Font size when displaying text.
text_italic	Boolean	*True* if the font is italic when displaying text.
text_bold	Boolean	*True* if the font is bold when displaying text.
feature_size	Integer	Font size for interlinear text.
feature_italic	Boolean	*True* if the font is italic for interlinear text.
feature_bold	Boolean	*True* if the font is bold for interlinear text.
tooltip_size	Integer	Font size for text in the grammar information box.
tooltip_italic	Boolean	*True* if the font is italic for text in the grammar information box.
tooltip_bold	Boolean	*True* if the font is bold for text in the grammar information box.
input_size	Integer	Font size in input fields.

(Continued...)

Column	Type	Contents
input_italic	Boolean	*True* if the font is italic in input fields.
input_bold	Boolean	*True* if the font is bold in input fields.

17.7 The *personal_font* Table

Each user can specify one personal font per alphabet. The personal font is listed on the font selection page together with the system fonts.

The *personal_font* table contains a user's personal fonts for a specific alphabet. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
user_id	Integer	The *id* field from an entry in the *user* table.
alphabet_id	Integer	The *id* field from an entry in the *alphabet* table.
font_family	Text	The name of the font.

17.8 The *exercisedir* and *classexercise* Tables

Together, the *exercisedir* and *classexercise* tables control which classes have access to which exercise directories. The *exercisedir* assigns an ID to each exercise directory. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
pathname	Text	The pathname of a directory, relative to the `quizzes` diretory.

The *classexercise* has an entry for each class that is allowed to access a given directory. It has the following fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
classid	Integer	The *id* field from an entry in the *class* table. Users belonging to this class have access to the exercises in the directory identified by the field *pathid*. If the *classid* field is 0, everybody has access to the exercises in the directory identified by the field *pathid*.
pathid	Integer	The *id* field from an entry in the *exercisedir* table.

17.9 The *bible_refs* Table

The *bible_refs* table contains links between Bible verses and pictures on the resource website (see Section 20.2) that are relevant for the verse. The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
book	Text	The name of the book. (This is the name used internally in the Emdros database.)
booknumber	Integer	The number of the book. This is the same information held in the *book* field. The file `myapp/controllers/ctrl_pic2db.php` contains an array that translates between book name and book number.
chapter	Integer	The chapter.
verse	Integer	The verse.
picture	Integer	The number of a picture on the resource website.

17.10 The *bible_urls* Table

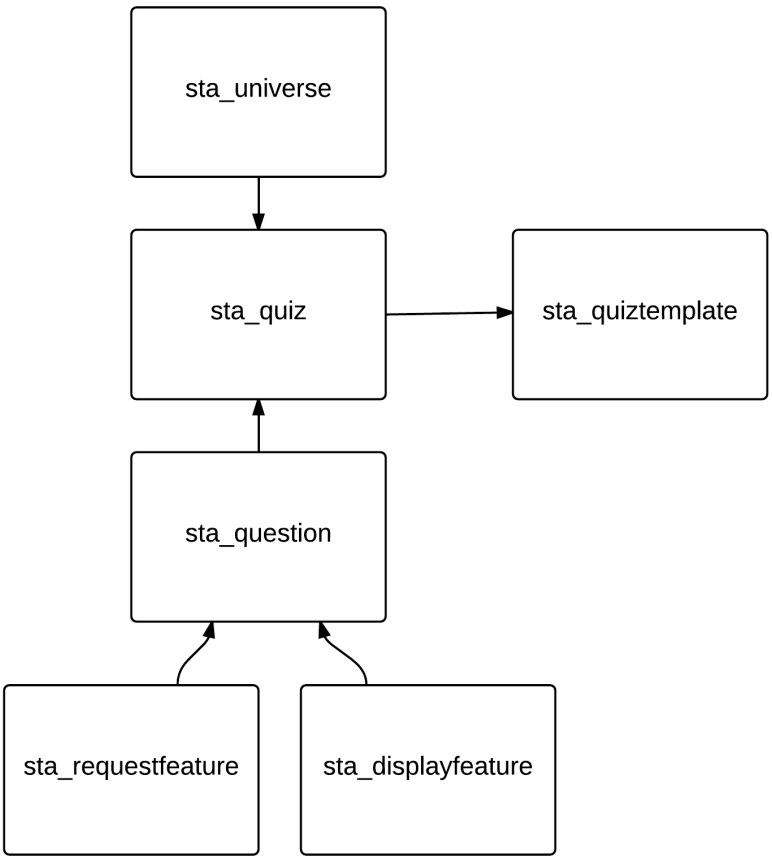
The *bible_urls* table contains links between Bible verses and URLs.[1] The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
book	Text	The name of the book. (This is the name used internally in the Emdros database.)
booknumber	Integer	The number of the book. This is the same information held in the *book* field. The file `myapp/controllers/ctrl_pic2db.php` contains an array that translates between book name and book number.
chapter	Integer	The chapter.
verse	Integer	The verse.
url	Text	The relevant URL.
type	Text	A single character identifying the type of icon to display. This character must be either D (for "Document"), V (for "Video"), or U (for "other URL").

17.11 The Statistics Tables

The six tables with names starting with *sta_* contain statistics about how well a user performed in an exercise. The following figure illustrates the relationship between the six tables.

[1] Currently, the URLs are configured on the resourse website (see Section 20.2), but they could come from other sources.



In this illustration, each arrow indicates a many-to-one relationship, with one item at the arrow head and many items at the other end of the arrow.

17.11.1 The *sta_quiz* Table

Every time a user starts running an exercise, an entry is created in the *sta_quiz* table. It contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this execution of an exercise.
userid	Integer	The *id* field from an entry in the *user* table. This identifies the user running the exercise.
templid	Integer	The *id* field from an entry in the *sta_quiztemplate* table. This identifies the quiz template used for this exercise.
start	Integer	The start time (UNIX time[2]).
end	Integer	The end time (UNIX time). This value is NULL if the exercise is still running or if the exercise was aborted without saving the result.
valid	Boolean	*False* if the user has deleted the entry, *true* otherwise. (There is currently no user interface for deleting statistics.

[2] That is, seconds since 00:00:00 UTC on 1 January 1970.

17.11.2 The *sta_quiztemplate* Table

Every time a user starts running an exercise, the system checks if the quiz template is already stored in the *sta_quiztemplate* table. If not, an entry is created. A template is identified by its contents; if, therefore, a facilitator changes the contents of a quiz template, a new entry will be created in the *sta_quiztemplate* table the next time the exercise is run.

The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this quiz template.
userid	Integer	The *id* field from an entry in the *user* table. This identifies the user running the exercise. For historical reasons, each user has their own set of templates in this table.
pathname	Text	The full pathname of the quiz template file.
dbname	Text	The name of the Emdros database on which the quiz template is based. This will normally be "ETCBC4" for the Hebrew Old Testament and "nestle1904" for the Greek New Testament.
dbproname	Text	The name of the localization information for the Emdros database on which the quiz template is based. This will normally be "ETCBC4" for the Hebrew Old Testament using the Hebrew alphabet, "ETCBC4-translit" for the Hebrew Old Testament using the transliterated alphabet, and "nestle1904" for the Greek New Testament.
qoname	Text	The Emdros type name of the sentence unit (quiz object) on which the exercise is based.
quizcode	Text	The actual XML text of the quiz template. Since this text contains both the database name, the database localization name, and the sentence unit, the table fields *dbname, dbpropname,* and *qoname* are actually superfluous, but they are included as separate fields to make decoding the XML text unnecessary in most cases.
quizcodehash	Integer	A hash value of the *quizcode* field. It can be used to speed up the comparison of the quizcode field from different entries in this table: If the *quizcodehash* values are different, then the *quizcode* values will also be different.

17.11.3 The *sta_universe* Table

Each entry in the *sta_universe* table represents a single book, chapter, or verse from the Bible. Together, a number of entries with the same *quizid* field identify the passages used for generating a particular exercise.

The *sta_universe* table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this entry.
userid	Integer	The *id* field from an entry in the *user* table.
quizid	Integer	The *id* field from an entry in the *sta_quiz* table.
component	Integer	A reference to a single book, chapter, or verse. The format is either "Genesis", "Genesis:3", or "Genesis:3:8".

17.11.4 The *sta_question* Table

Each entry in the *sta_question* table represents a single question, as defined in Section 1.2 (see Figure 1.1 on page 6). The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this question.
userid	Integer	The *id* field from an entry in the *user* table.
quizid	Integer	The *id* field from an entry in the *sta_quiz* table.
txt	Text	The text of the question. The quiz objects are enclosed between `<em>` and `</em>`.
location	Text	The Bible reference for the text, given in the format "Genesis, 3, 8".
time	Integer	The time (UNIX time) when this question was answered.

17.11.5 The *sta_displayfeature* Table

Each entry in the *sta_displayfeature* table lists a display feature that was shown for a question item (see Section 1.2 and Figure 1.1 on page 6 for a definition of "question item".)

The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this question.
userid	Integer	The *id* field from an entry in the *user* table.
questid	Integer	The *id* field from an entry in the *sta_question* table.
qono	Integer	The index (starting from 1) of the question item within the question.
name	Text	The name of the feature.
value	Text	The value of the feature.

17.11.6 The *sta_requestfeature* Table

Each entry in the *sta_requestfeature* table lists a request feature that was required for a question item (see Section 1.2 and Figure 1.1 on page 6 for a definition of "question item".)

The table contains these fields:

Column	Type	Contents
id	Integer	Unique number identifying this question.
userid	Integer	The *id* field from an entry in the *user* table.

(Continued...)

Column	Type	Contents
questid	Integer	The *id* field from an entry in the *sta_question* table.
qono	Integer	The index (starting from 1) of the question item within the question.
name	Text	The name of the feature.
value	Text	The correct value of the feature.
answer	Text	The answer provided by the user. If the user tries to answer several times in a single exercise only the first answer is recorded.
correct	Boolean	*True* if the user's answer is correct.[3]

17.12 Initial Database Setup

Before you run a Bible OL system for the first time, the database must initialized to sensible values. The file `bolsetup.sql` contains SQL commands that do this.

Additionally you must manually add a user to the *user* table with administrative rights. The entry should set the field *isadmin* to *true* (1). The *password* should be set to the value of this PHP function:

```
md5($config['pw_salt'] . 'xxxxx');
```

where *$config['pw_salt']* is configured in `myapp/config/ol.php`, and *xxxxx* is the user's desired password.

[3] This is not the same as testing if *value=answer*. For example, when providing an English translation of a word, a correct *answer* may be "do" even if the *value* is "make, do; fix; deal with".

Chapter 18

Less Style Sheets

Read this chapter if you are going to understand or modify style sheets.

CSS style sheets can sometimes be unwieldy to work with. The *Less* program is a CSS preprocessor that allows a clearer way to structure style sheets.

Compare, for example, the CSS code in the left column below with the Less code in the right column:

CSS

```
ul.dropdown {
  padding: 0;
}

ul.dropdown a {
  text-decoration: none;
}

ul.dropdown li {
  display: inline-block;
  background: #f3d673;
  z-index: 1;
}

ul.dropdown li:hover {
  background: #c0c0c0;
  position: relative;
}

ul.dropdown li a {
  color: black;
  display: block;
}
```

Less

```
@beige: #f3d673;
@mediumgray: #c0c0c0;

ul.dropdown {
    padding: 0;

    a {
        text-decoration: none;
    }

    li {
        display: inline-block;
        background: @beige;
        z-index: 1;

        &:hover {
            background: @mediumgray;
            position: relative;
        }

        a {
            color: black;
            display: block;
        }
    }
}
```

Section 6.2 describes how to install the Less compiler *lessc*. Details of the Less language can be found at `http://lesscss.org`.

Although Less style files can be compiled when used in a browser, the Bible OL implementation compiles Less files only once and stores the resulting CSS files. This is achieved through the Makefile in the top directory. The command "make styles/ol.css" will compile the Less file.[1]

[1] The simple command "make" will compile all Less and TypeScript files.

At present, Bible OL uses only one Less file, namely `styles/ol.less` which compiles into `styles/ol.css`.

Chapter 19

Client Code

Read this chapter if you are going to understand or modify the client code.

The client code runs in a web browser. Most of it is written in TypeScript which is compiled into JavaScript.

There are three different TypeScript programs that can run as client code:

- *ol,* which displays text or runs exercises. (See Section 19.2.)
- *editquiz,* which edits a quiz template.
- *fontselector,* which allows a user to set font preferences.

19.1 TypeScript

TypeScript is a superset of JavaScript that adds strong typing and proper classes to JavaScript. The website `http://www.typescriptlang.org` contains a tutorial and the formal specification of the language.

Section describes how to install the TypeScript compiler *tsc.*

The Bible OL implementation compiles TypeScript files only once and stores the resulting JavaScript files. This is achieved through the Makefile in the top directory. The command "make all" (or simply "make") will compile the TypeScript files (and the Less file). The Makefile checks that the *tsc* compiler version is 1.0.1.0. This is probably unnecessary, and should be removed as newer versions appear.

The TypeScript files are found in the directory `ts`; the resulting JavaScript files are stored in the directory `js`.

19.2 The *ol* Client Code

The *ol* client code is responsible for displaying text and running an exercise based on information provided by the server – primarily in the JavaScript variables *configuration, localization, typeinfo, dictionaries,* and *quizdata,* which are described in detail in Chapters 12 and 13.

For text display, the *ol* program buids the text inside an HTML skeleton provided by the server. The skeleton looks like this (somewhat simplified):

```
<div class="grammarselector" id="gramselect"></div>
<div class="grammardisplay"></div>

<div id="textcontainer">
    <h1></h1>

    <div id="textarea"></div>

    <p><button id="togglemql">Toggle MQL</button></p>
```

```
    <pre class="mqlarea">SELECT ALL OBJECTS WHERE ...</pre>
</div>
```

When running an exercise, the *ol* program buids a question inside an HTML skeleton provided by the server. The skeleton looks like this (somewhat simplified):

```
<div class="grammarselector" id="gramselect"></div>
<div class="grammardisplay"></div>

<div id="textcontainer">
    <div id="quizdesc"></div>

    <div id="textarea"></div>

    <div id="virtualkbcontainer"><div id="virtualkbid"></div></div>

    <table id="quiztab"></table>

    <p><input id="locate_cb" type="checkbox">Locate: <span class="location"></span></p>

    <p class="inline">Progess:</p>
    <progress id="progress" value="0" max="1"></progress>
    <div id="progressbar"></div>
    <p id="progresstext" class="inline"></p>

    <div id="buttonlist1">
        <button id="check_answer" type="button">Check answer</button>
        <button id="show_answer" type="button">Show answer</button>
    </div>
    <div id="buttonlist2">
        <button id="next_question" type="button">Next</button>
        <button id="finish" type="button">Finish</button>
    </div>

    <p><button id="togglemql">Toggle MQL</button></p>
    <pre class="mqlarea">SELECT ALL OBJECTS WHERE ...</pre>
</div>
```

The `<div>` areas of class *grammarselector* and *grammardisplay* are for the grammar selection box and grammar information box, respectively. The *ol* program builds the contents of based. The *grammarselector* is built using the class *GenerateCheckboxes*; the *grammardisplay* is built by the function *toolTipFunc*, which is defined within the function *generateSentenceHtml* in the *Dictionary* class.

The `<div>` area of class *textcontainer* contains the text and, possibly, the question. The actual text is places in the `<div>` area of class *textarea*. The question items are placed in the `<table>` with *id="quiztab"*.

The `<button>` with *id="togglemql"* and the `<pre>` area of class *mqlarea* are normally not shown to the user. They are intended for debugging only. The *mqlarea* contains the MQL commands executed during the creation of the text or exercise. You can either inspect the *mqlarea* by looking at the HTML source sent to the browser, or you can enable the `<button>` by removing "`display: none`" from this instruction in `styles/ol.less`:[1]

```
button#togglemql {
    display: none;
}
```

When the "`display: none`" line has been removed, a "Toggle MQL" button will appear in the browser. Clicking the button will display the MQL commands executed during the creation of the text

[1] After modifying `styles/ol.less`, you must recompile the Less file. For simple debugging, you may prefer to edit the `styles/ol.css` file directly.

or exercise.

The most complicated task for *ol* is probably to build the contents of the *textarea*, and this will be described in some detail below.

As Chapter 13 explains, the *dictionaries* variable contains the field *sentenceSets,* which is an array of *MonadSet* objects, and the field *monadObjects,* which is an array of arrays of arrays of *MonadObject* objects. When Bible OL is displaying text, the array *sentenceSets* and the top array in *monadObjects* have only one element; but when Bible OL is displaying an exercise consisting of n questions, the two arrays have n elements.

The *ol* program converts the *dictionaries* variable (which is of interface class *DictionaryIf*) into one or more objects of class *Dictionary*, one for each entry in the *MonadSets/MonadObject* arrays. (Note the *Dictionary* class here must not be confused with the *Dictionary* class in the server. The server's *Dictionary* class corresponds to the client's *DictionaryIf* interface.)

As explained in Section 13.1, the *monadObjects* field of the *DictionaryIf* interface is an array of array of arrays. The middle array is indexed by the level in the grammatical hierarchy (word, phrase, clause, etc.). As part of creating a *Dictionary* from a given index in a *DictionaryIf*, the *constructor* function in the *Dictionary* class builds a parallel collection of arrays: Each *MonadObjects* is complemented by one or more *DisplayMonadObjects.* A *DisplayMonadObject* represents the physical appearance of an Emdros object in the browser. *DisplayMonadObject* has a member function, *generateHtml,* which is responsible for generating the HTML that renders the Emdros object.

Just as a *MonadObject* is either a *SingleMonadObject* or a *MultipleMonadObject*, a *DisplayMonadObject* is either a *DisplaySingleMonadObject* (typically representing a word) or a *DisplayMultipleMonadObject* (typically representing a phrase, clause, or sentence). There is, however, an important difference between a *MultipleMonadObject* and a *DisplayMultipleMonadObject.* If, for example, a clause consists of multiple noncontiguous parts, it is represented by one *MultipleMonadObject* but by multiple *DisplayMultipleMonadObjects,* one for each part of the clause.

When the *generateHtml* is called for a *DisplaySingleMonadObject,* its task is to create HTML code to represent a single word and all its features. The code generated is structured like this:

```
<span class="textblock inline">
    <span class="textdisplay charset" data-idd="ID_D">text</span>

    <span class="wordgrammar dontshowit featurename charset">featurevalue</span>
    <span class="wordgrammar dontshowit featurename charset">featurevalue</span>
    <span class="wordgrammar dontshowit featurename charset">featurevalue</span>
    ...

</span>
```

Here,

charset	identifies the character set and hence the font and text direction. Valid values are `hebrew`, `hebrew_translit`, `greek`, `latin`, and `ltr`. The value `ltr` is used to force left-to-right for features in Latin script. The value `latin` is currently not used; it is reserved for corpuses that use the Latin alphabet.
ID_D	is the ID_D (see Chapter 5) of the word.
text	is the actual word.
featurename	is the non-localized name of the feature, that is, the name of the feature as it appears in the Emdros database.
featurevalue	is the localized value of the feature.

If a particular feature is turned on in the grammar selection box, the `dontshowit` class in the relevant `<span>` elements is changed to `showit`.

Two extra class values are added to the `<span class="textdisplay">` element to control the rendering of Hebrew word spacing.

The first class value is one of these and doesn't change:

Class value	Meaning
`cont`	The word must be followed immediately by the next word with no intervening space.
`contx`	The word ends in a *maqaf* (־) and must be followed immediately by the next word with no intervening space.

If neither `cont` nor `contx` is set on a word, a `<span>` element with class `wordspace` is inserted after the word.

The second class value is one of these, and it changes as the user switches between display and not displaying word spacing:

Class value	Meaning
`cont1`	The user has not requested word spacing. Use default rending of words.
`cont2`	The user has requested word spacing. Add a hyphen and a space to the end of the current word.
`cont2x`	The user has requested word spacing. The current word ends in a *maqaf*. Add a space to the end of the current word.

When the *generateHtml* is called for a *DisplayMultipleMonadObject*, its task is to create HTML code to represent a single word and all its features. The code generated is structured like this:

```
<span class="notdummy nolevel noseplin">
    <span class="gram dontshowit" data-idd="ID_D">loctype
        <span class="xgrammar dontshowit type_featurename">:featurevalue</span>
        <span class="xgrammar dontshowit type_featurename">:featurevalue</span>
        <span class="xgrammar dontshowit type_featurename">:featurevalue</span>
        ....
    </span>
    ... (Lower levels in the grammar hierarchy are inserted here)
</span>
```

Here,

nolevel	is one of `nolev1`, `nolev2`, `nolev3`, etc. The number within this name identifies the level in the grammar hierarchy: 1 for the level just above *word*, 2 for the next higher level, etc.
ID_D	is the ID_D (see Chapter 5) of the word.
loctype	is the localized name of the Emdros object type.
type	is the non-localized name of the Emdros object type, that is, the name of the type as it appears in the Emdros database.
featurename	is the non-localized name of the feature, that is, the name of the feature as it appears in the Emdros database. Note that *type* and *featurename* are strung together with an intervening underscore, thus forming a single class value.
featurevalue	is the localized value of the feature.

The `<span class="notdummy">` element may additionally have the class value `hasp` and/or `hass`. This indicates that the current *DisplayMultipleMonadObject* is part of a noncontiguous collection of monads. The class value `hasp` means that the *DisplayMultipleMonadObject* has a predecessor; the class value `hass` means that the *DisplayMultipleMonadObject* has a successor.

If "Show border" is selected in the grammar selection box for a particular level in the grammar hierarchy, the `dontshowit` class in the relevant `<span class="gram">` elements is changed to `showit`.

If "Seperate lines" is selected in the grammar selection box for a particular level in the grammar hierarchy, the `noseplin` class in the relevant `<span class="gram">` elements is changed to `seplin`.

If a particular feature is turned on in the grammar selection box, the `dontshowit` class in the relevant `<span class="xgrammar">` elements is changed to `showit`.

Occasionally[2], a level in the grammar hierarchy is missing. If this is the case, the server code will insert a dummy object in the hierarchy, and the client will generated this HTML code:

```
<span class="nolevel noseplin">
    <span class="nogram dontshowit" data-idd="ID_D">loctype
    </span>
    ... (Lower levels in the grammar hierarchy are inserted here)
</span>
```

Note the absense of the `notdummy` class value.

At the top (patriarch) level the `notdummy` class name is omitted, and at this level the HTML simply looks like this:

```
<span class="nolevel noseplin">
    ... (Lower levels in the grammar hierarchy are inserted here)
</span>
```

Note the absense of the `notdummy` class value.

[2] Currently only in the neste1904 database.

Chapter 20

Complementary Websites

Read this chapter if you want to.

A few addtional websites complement the function of Bible OL: The Learning Journey website, the resource website, and the SHEBANQ website.

20.1 The Learning Journey Website

Bible OL collects information about the answers users give to exercises and how long time they spend on each exercise. The Learning Journey access this information and provides statistics about each user's performance. Currentely, the Learning Journey URL is `http://statdb.3bmoodle.dk`.

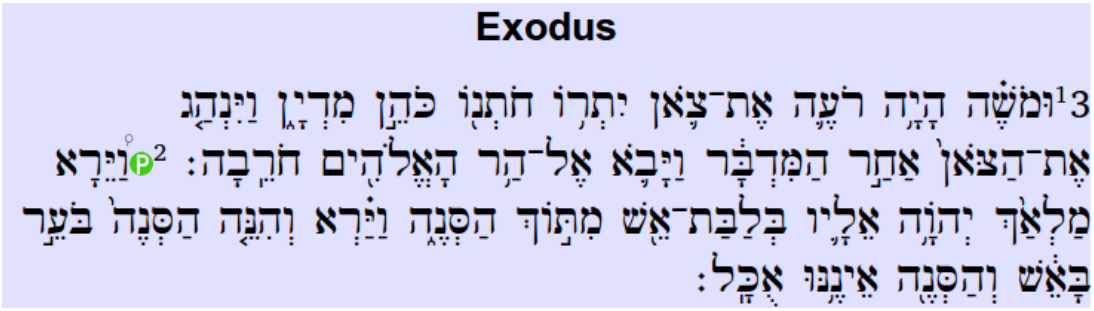
Learning Jourey directly accesses the statitics and user tables of the Bible OL user database (See Chapter 17.

No further information about Learning Journey is given in this document.

20.2 The Resource Website

The resource web site is a collection of photos from the Middle East. Many of them relate to events and places described in the Bible. The photos have descriptive texts that contain Bible references. The URL of the resource website is `http://resources.3bmoodle.dk`.

Bible OL can use information from the resource website to add picture links to Bible passages. If a photo in the resource website refers to, for example, Exodus 3 : 2, and a user ticks the "Show link icons" checkbox when displaying Exodus chapter 3, a green "P" icon will appear in the text next to verse 2:



Clicking on the icon will cause the web browser to display the relevant photo. If there are more than one photo, the icon will be blue rather than green.

Bible OL gathers information about the photos on the resource web site by executing the command

```
php index.php pic2db
```

in a cron job. This command calls Bible OL's *Ctrl_pic2db* controller (in the file `myapp/controllers/ctrl_pic2db.php`), which requests information from the resource database.

By accessing the URL `http://resources.3bmoodle.dk/jsonrefs.php`, the *Ctrl_pic2db* controller receives a JSON object from the resources website containing information about the photos and the Bible verses to which they refer. Bible OL stores this information in the *bible_refs* table in the user database (see Section 17.9).

In addition to the pictures, the resource website may also provide URLs associated with various Bible verses. These URLs are configured in the resource website but are otherwise unrelated to the functioning of that website. The URLs are intended to identify videos, documents, or other resources that may relevant for studying a particular verse.

Information about these URLs is also retrieved by the cron job above. By accessing the URL `http://resources.3bmoodle.dk/jsonurls.php`, the *Ctrl_pic2db* controller receives a JSON object from the resources website containing information about the URLs and the Bible verses to which they refer. Bible OL stores this information in the *bible_urls* table in the user database (see Section 17.10). Links to these urls are displayed as "V", "D", or "U" icons.

20.3 The SHEBANQ Website

SHEBANQ (System for HEBrew text: ANnotations for Queries and markup) is a website that uses the ETCBC4 database for displaying text and grammar information for the Hebrew Bible. The URL is `http://shebanq.ancient-data.org`.

When Bible OL displays a text from the Old Testament, an icon in the upper right corner of the text area provides a link to the same chapter at the SHEBANQ website. A similar link to Bible OL is found on the SHEBANQ website.

Appendix A

ETCBC4 Details

The *ETCBC4* Emdros database contains the Hebrew and Aramaic text for the Old Testament.

The database comes from the *Eep Talstra Center for Bible and Computer* and is made available under a Creative Commons Attribution-NonCommercial 4.0 International License.[1] When describing the database, a text similar to this one should be used: "The database itself can be found through this persistent identifier: `urn:nbn:nl:ui:13-048i-71`." The identifier should be a hyperlink pointing to `http://www.persistent-identifier.nl/?identifier=urn:nbn:nl:ui:13-048i-71`.

Prior to using this database in Bible OL, I have added additional features to the information from its original creators. More details about this is given in Section A.1.21.

A.1 The *word* Object

The basic object type is the *word*. Each *word* corresponds to a single monad. The following sections give details about the features of the object.

Many of the features exist in several different encodings. The encodings are indicated by the name of the feature. A feature *XXX* may exists in these variants:

Feature name	Encoding
XXX	Transcribed alphabet
XXX_utf8	Native alphabet
XXX_translit	Transliterated alphabet
XXX_cons_utf8	Consonants only, native alphabet
XXX_nocant_utf8	Cantillation marks omitted, native alphabet
XXX_nopunct_translit	Punctuation omitted, transliterated alphabet

Except where otherwise noted, all features of the *word* object are string features.

In cases where a feature is an enumeration, the enumeration type may contain values that are not actually used. Such values are not listed in the following sections.

A.1.1 Features: text, continuation, g_word, suffix, and g_cons

These feature relate to the visual appearance of a word.

The *text* features contain the actual text of the word. These variants exist:

- text
- text_utf8
- text_translit
- text_cons_utf8
- text_nocant_utf8
- text_nopunct_translit

[1] `http://creativecommons.org/licenses/by-nc/4.0`.

The *text* features work together with the *continuation* features as described in Section 8.2.2. The *continuation* features exist in these variants:

- continuation
- continuation_utf8
- continuation_translit

The *g_word* features are very similar to the *text* features. The *g_word* feature is not used by Bible OL because its information is available by other means. The difference between *g_word* and *text* is that *g_word* does not use the proper final form of the Hebrew consonants, and it includes punctuation information that follows a word. This punctuation information is typically the character ׃ (sof pasuq) or ־ (maqaf), but it may also include additional characters.

The *g_word* features exist in these variants:

- g_word
- g_word_utf8
- g_word_cons_utf8

The *suffix* features contain characters that follow a word (as described above for *g_word*). Possible suffixes are:

- ־
- ׃
- ם
- ף
- ן׃
- ם ן׃
- ף ן׃
- ם׃
- ף׃

The *suffix* features exist in these variants:

- suffix
- suffix_translit
- suffix_utf8

Note: Do not confuse these suffix features with the features *suffix_gender, suffix_number,* and *suffix_person* which are described in Section A.1.14.

The *g_cons* features contain the consonants of the word. They are not used by Bible OL because their information is available by other means. The *g_cons_utf8* feature is identical to the *text_cons_utf8* feature, except that *g_cons_utf8* does not use the proper final form of the Hebrew consonants.

The *g_cons* features exist in these variants:

- g_cons
- g_cons_utf8

A.1.2 Features: g_lex, lex, and vocalized_lexeme

The "lexeme" or "lemma" of a word is the version of the word found in a dictionary. For example, the English word "mice" has the lexeme "mouse" because in a dictionary, the word is found under the entry "mouse".

ETCBC4 has three different lexeme features with different characteristics.

The *vocalized_lexeme* is the word commonly taken to be the lexeme of a Hebrew word. Except for verbs, this lexeme contains vowels. The *vocalized_lexeme* features exist in these variants:

- vocalized_lexeme
- vocalized_lexeme_utf8

- vocalized_lexeme_translit
- vocalized_lexeme_cons_utf8

The *g_lex* and *lex* features contain various other ways to write the lexeme. Compare, for example, these three lexemes for אֱלֹהִים:

Feature name	Content
vocalized_lexeme_utf8	אֱלֹהִים
g_lex_utf8	אֱלֹה
lex_utf8	אלהימ (Note the shape of the מ)

I have not studied all the differences between the three lexemes. Suffice it to say that *vocalized_lexeme_utf8* is the one normally shown to users, and *lex* is useful internally in the system because it only contains consonants and uses the transcribed alphabet.

The *g_lex* features exist in these variants:

- g_lex
- g_lex_utf8
- g_lex_cons_utf8

The *lex* features exist in these variants:

- lex
- lex_utf8
- lex_cons_utf8

A.1.3 Features: lexeme_occurrences and frequency_rank

ETCBC4 contains information about the frequency of various lexemes. The feature *lexeme_occurrences* is an integer feature containing the number of times this lexeme occurs in the Old Testament. The feature *frequency_rank* is an integer feature containing the rank of each lexeme: The most frequent lexeme has rank 1, the second most frequent lexeme has rank 2, etc.

A.1.4 Features: english and german

The features *english* and *german* give English and German glosses for the lexeme. The string may contain several meanings separated by commas and semicolons. Also, special information may be given between < and > characters. For example:

Lexeme	English gloss
רֵאשִׁית	beginning; what is first; best
אֵת	<object marker>

A.1.5 Features: pfm, g_pfm

The *pfm* feature describes the paradigmatic form of the preformative. The following values are possible:

- The empty string
- >
- absent[2]
- H
- J
- L

[2] This is the actual string "absent".

- M
- N
- n/a
- T=
- T

The *g_pfm* features contain the graphical representation of the preformative. These features exist in these variants:

- g_pfm
- g_pfm_utf8
- g_pfm_translit
- g_pfm_cons_utf8

A.1.6 Features: vbs, g_vbs

The *vbs* feature describes the paradigmatic form of the root formation morpheme. The following values are possible:

- >
- absent[3]
- C
- H
- HCT
- HT
- N
- n/a
- NT
- >T
- T

The *g_vbs* features contain the graphical representation of the root formation morpheme. These features exist in these variants:

- g_vbs
- g_vbs_utf8
- g_vbs_translit
- g_vbs_cons_utf8

A.1.7 Features: vbe, g_vbe

The *vbe* feature describes the paradigmatic form of the verbal ending. The following values are possible:

- The empty string
- H=
- H
- J
- JN
- N>
- N
- n/a
- NH
- NW
- T==

[3] This is the actual string "absent".

- T=
- T
- TJ
- TM
- TN
- TWN
- W
- WN

The *g_vbe* features contain the graphical representation of the verbal ending. These features exist in these variants:

- g_vbe
- g_vbe_utf8
- g_vbe_translit
- g_vbe_cons_utf8

A.1.8 Features: nme, g_nme

The *nme* feature describes the paradigmatic form of the nominal ending. The following values are possible:

- The empty string
- absent[4]
- H
- J=
- J
- JM=
- JM
- JN=
- JN
- N
- n/a
- T=
- T
- TJ
- TJM
- TJN
- W=
- W
- WT
- WTJ

The *g_nme* features contain the graphical representation of the nominal ending. These features exist in these variants:

- g_nme
- g_nme_utf8
- g_nme_translit
- g_nme_cons_utf8

[4]This is the actual string “absent”.

A.1.9 Features: uvf, g_uvf

The *uvf* feature describes the paradigmatic form of the univalent final. The following values are possible:

- >
- absent[5]
- H
- J
- N
- W

The *g_uvf* features contain the graphical representation of the univalent final. These features exist in these variants:

- g_uvf
- g_uvf_utf8
- g_uvf_translit
- g_uvf_cons_utf8

A.1.10 Features: prs, g_prs

The *prs* feature describes the paradigmatic form of the pronominal suffix. The following values are possible:

	Meaning		
Value	**Gender**	**Person**	**Number**
absent[6]		Absent	
H	Feminine	Third	Singular
H=	Masculine	Third	Singular
HJ	Masculine	Third	Singular
HM	Masculine	Third	Plural
HN	Feminine	Third	Plural
HW	Masculine	Third	Singular
HWN	Masculine	Third	Plural
J	Common	First	Singular
K	Masculine	Second	Singular
K=	Feminine	Second	Singular
KM	Masculine	Second	Plural
KN	Feminine	Second	Plural
KWN	Masculine	Second	Plural
M	Masculine	Third	Plural
MW	Masculine	Third	Plural
N	Feminine	Third	Plural
N>	Common	First	Plural
NJ	Common	First	Singular
NW	Common	First	Plural
W	Masculine	Third	Singular
n/a		Not applicable	

The *g_prs* features contain the graphical representation of the pronominal suffix. These features exist in these variants:

[5] This is the actual string "absent".
[6] This is the actual string "absent".

- g_prs
- g_prs_utf8
- g_prs_translit
- g_prs_cons_utf8

A.1.11 Feature: g_qere

The *g_qere* features contain the qere reading for the current word, if any.

The *g_qere* features exist in these variants:

- g_qere
- g_qere_utf8
- g_qere_translit

A.1.12 Feature: language

The *language* feature indicates if the word is Hebrew or Aramaic. It is an enumeration feature of type *language_t* whose value is either *Hebrew* or *Aramaic.*

A.1.13 Features: sp and pdp

The *sp* feature indicates the part of speech of the word; the *pdp* feature indicates the phrase dependent part of speech. Both are enumeration features of type *part_of_speech_t.* They can have these values:

Value	Meaning
adjv	Adjective
advb	Adverb
art	Article
conj	Conjunction
inrg	Interrogative
intj	Interjection
nega	Negative
nmpr	Proper noun
prde	Demonstrative pronoun
prep	Preposition
prin	Interrogative pronoun
prps	Personal pronoun
subs	Noun
verb	Verb

A.1.14 Features: ps, nu, gn, suffix_person, suffix_number, suffix_gender

The *ps, nu,* and *gn* features indicate the person, number, and gender, respectively, of the word. They are enumeration features of type *person_t, number_t,* and *gender_t* respectively.

The *suffix_person, suffix_number,* and *suffix_gender* features indicate the person, number, and gender, respectively, of an optional suffix on the word. (These values are derived from the *prs* feature described in Section A.1.10.) They are enumeration features of type *suffix_person_t, suffix_number_t,* and *suffix_gender_t* respectively. Note: Do not confuse these features with the suffix features described in Section A.1.1.

The *person_t* and *suffix_person_t* enumerations have these values:

Value	Meaning
p1	First person
p2	Second person
p3	Third person
unknown	Unknown person (only legal for *person_t*)
NA	Not applicable

The *number_t* and *suffix_number_t* enumerations have these values:

Value	Meaning
sg	Singular
du	Dual (only legal for *number_t*)
pl	Plural
unknown	Unknown number (only legal for *number_t*)
NA	Not applicable

The *gender_t* and *suffix_gender_t* enumerations have these values:

Value	Meaning
f	Feminine
m	Masculine
c	Common (only legal for *suffix_gender_t*)
unknown	Unknown number (only legal for *gender_t*)
NA	Not applicable

A.1.15 Feature: ls

The *ls* feature indicates the lexical set of the word. It is an enumeration feature of type *lexical_set_t* whose value is one of the following:

Value	Meaning
afad	Anaphoric adverb
card	Cardinal
cjad	Conjunctive adverb
focp	Focus particle
gntl	Gentilic
mult	Noun of multitude
nmcp	Copulative noun
nmdi	Distributive noun
none	None
ordn	Ordinal
padv	Potential adverb
ppre	Potential preposition
ques	Interrogative particle
quot	Quotation verb
vbcp	Copulative verb

A.1.16 Feature: vs

The *vs* feature indicates the verbal stem of the word. It is an enumeration feature of type *verbal_stem_t* whose value is one of the following:

Value	Meaning	Used in language
afel	Afel	Aramaic
etpa	Etpaal	Both
etpe	Etpeel	Aramaic
haf	Hafel	Aramaic
hif	Hifil	Hebrew
hit	Hitpael	Hebrew
hof	Hofal	Both
hotp	Hotpaal	Hebrew
hsht	Hishtafal	Both
htpa	Hitpaal	Aramaic
htpe	Hitpeel	Aramaic
nif	Nifal	Hebrew
nit	Nitpael	Hebrew
pael	Pael	Aramaic
pasq	Passive Qal	Hebrew
peal	Peal	Aramaic
peil	Peil	Aramaic
piel	Piel	Hebrew
pual	Pual	Hebrew
qal	Qal	Hebrew
shaf	Shafel	Aramaic
tif	Tifal	Hebrew
NA	Not applicable	

A.1.17 Feature: vt

The *vt* feature indicates the verbal tense of the word. It is an enumeration feature of type *verbal_tense_t* whose value is one of the following:

Value	Meaning
coho	Cohortative
emim	Emphatic imperative
impf	Imperfect
impv	Imperative
infa	Infinitive absolute
infc	Infinitive construct
juss	Jussive
perf	Perfect
ptca	Participle
ptcp	Passive participle
wayq	Wayyiqtol
NA	Not applicable

A.1.18 Feature: st

The *st* feature indicates the state of the word. It is an enumeration feature of type *state_t* whose value is one of the following:

Value	Meaning
a	Absolute
c	Construct
e	Emphatic
NA	Not applicable

A.1.19 Feature: verb_class

The *verb_class* feature indicates the verb classes to which a word belongs. It is list of values of the enumeration type *verb_class_t* whose values are:

- analog_i_nun
- analog_i_waw
- four_consonants
- geminate
- hjh_xjh
- i_aleph
- i_guttural
- i_nun
- i_waw
- i_yod
- ii_guttural
- ii_waw
- ii_yod
- iii_aleph
- iii_guttural
- iii_he
- regular

A.1.20 Features: number, distributional_parent, functional_parent

These features are currently not used by Bible OL and are not documented here.

A.1.21 The Origin of the Features

The following features of the *word* object were part of the ETCBC4 database as I received it from its creators:

distributional_parent	g_uvf_utf8	nu
functional_parent	g_vbe	number
g_cons	g_vbe_utf8	pdp
g_cons_utf8	g_vbs	pfm
g_lex	g_vbs_utf8	prs
g_lex_utf8	g_word	ps
g_nme	g_word_utf8	sp
g_nme_utf8	gn	st
g_pfm	language	uvf
g_pfm_utf8	lex	vbe
g_prs	lex_utf8	vbs
g_prs_utf8	ls	vs
g_uvf	nme	vt

The following word features were generated by my program *emdros_updater* which is currently not published or documented:

continuation	g_qere_utf8	suffix_person
continuation_translit	g_uvf_cons_utf8	suffix_translit
continuation_utf8	g_uvf_translit	suffix_utf8
english	g_vbe_cons_utf8	text
frequency_rank	g_vbe_translit	text_cons_utf8
g_lex_cons_utf8	g_vbs_cons_utf8	text_nocant_utf8
g_nme_cons_utf8	g_vbs_translit	text_nopunct_translit
g_nme_translit	g_word_cons_utf8	text_translit
g_pfm_cons_utf8	german	text_utf8
g_pfm_translit	lex_cons_utf8	verb_class
g_prs_cons_utf8	lexeme_occurrences	vocalized_lexeme
g_prs_translit	suffix	vocalized_lexeme_cons_utf8
g_qere	suffix_gender	vocalized_lexeme_translit
g_qere_translit	suffix_number	vocalized_lexeme_utf8

The *emdros_updater* also fixed final versions of the characters kaph, mem, nun, pe, and tsade in the features *g_nme_utf8, g_prs_utf8,* and *g_vbe_utf8.*

Based on information from Assistant Professor Oliver Glanz of Andrews University, the verbal tenses *jussive, cohortative,* and *emphatic imperative* were added by a couple of (unpublished and undocumented) script in the *extra_tenses* directory under *emdros_updater.*

A.2 The Other Object Types

For information about the other object types in the ETCBC4 database, please consult the MQL code used for generating the database. The MQL code is found in the file TBD.

Appendix B

ETCBC4 Words Database

The Words Database (see Section 11) for the ETCBC4 Emdros database has the structure defined by these SQL statements:

```
CREATE TABLE lexemes (id integer primary key, lex text);
CREATE TABLE lexsuf (id integer primary key, lexid integer, sufid integer);
CREATE TABLE lextext (id integer primary key, lexid integer, textid integer);
CREATE TABLE suffixes (id integer primary key, suffix blob, suffix_translit blot);
CREATE TABLE texts (id integer primary key, word blob, word_translit blob);

CREATE INDEX ixlexemes on lexemes(lex);
CREATE INDEX ixlexsuf on lexsuf(lexid);
CREATE INDEX ixlextext on lextext(lexid);
CREATE INDEX ixsuffixes on suffixes(suffix);
CREATE INDEX ixsuffixes2 on suffixes(suffix_translit);
CREATE INDEX ixtexts on texts(word);
CREATE INDEX ixtexts2 on texts(word_translit);
```

The *lexemes* table has an entry for every possible value of the *lex* feature in the Hebrew parts (that is, not the Aramaic parts) of the Old Testament. Each entry consists of an ID number and the *lex* value.

The *suffixes* table has an entry for every possible pair of values of the *g_prs_utf8* and *g_prs_translit* features in the Hebrew parts of the Old Testament. Each entry consists of an ID number and the values of the *g_prs_utf8* and the *g_prs_translit* features.

The *texts* table has an entry for every possible pair of values of the *text_nocant_utf8* and *text_nopunct_translit* features in the Hebrew parts of the Old Testament. Each entry consists of an ID number and the values of the *text_nocant_utf8* and the *text_nopunct_translit* features.

The *lexsuf* table combines entries in the *lexemes* table with entries in the *suffixes* table. Similarly, the *lextext* table combines entries in the *lexemes* table with entries in the *texts* table. Thus, for example, the SQL statement

```
SELECT lex,suffix FROM lexemes
    JOIN lexsuf ON lexsuf.lexid=lexemes.id
    JOIN suffixes ON lexsuf.sufid=suffixes.id;
```

will list all possible combinations of the *lex* and *g_prs_utf8* features.

B.1 The Origin of the ETCBC4 Words Database

The current Word Database was originally generated for the old WIVU Emdros database. It was created using a Java program called `WordDb.java`, which has not been updated for ETCBC4.

Appendix C

Nestle1904 Details

The *nestle1904* database is in the public domain and derives from the 1904 version of Nestle's Greek New Testament text.

C.1 The *word* Object

The basic object type is the *word*. Each *word* corresponds to a single monad. The following sections give details about the features of the object.

Except where otherwise noted, all features of the *word* object are string features. They contain Unicode characters in UTF-8 encoding.

C.1.1 Features: surface, normalized, raw_normalized

The *surface* feature contains the actual text of the word.

The *normalized* feature is an attempt at a "normalized" form of the word. "Normalized" here means:

a) Punctuation has been removed.
b) Most accents due to throwback clitics have been eliminated.
c) Any final grave accent has been made acute when not eliminated by (b).

Note that process (b) is not perfect. It only normalizes words which have more than one accent. A consequence of this is that clitics such as μου will not get the accent removed even when the accent is present (e.g., due to a throwback clitic that follows it). Thus the *normalized* feature is not totally reliable.

The *raw_normalized* feature is the *normalized* feature with non-letter characters removed and all characters converted to lower case characters without accents.

C.1.2 Features: lemma and raw_lemma

The "lexeme" or "lemma" of a word is the version of the word found in a dictionary. For example, the English word "mice" has the lexeme "mouse" because in a dictionary, the word is found under the entry "mouse".

The *lemma* feature contains the lemma of the word. The *raw_lemma* feature is the lemma with non-letter characters removed and all characters converted to lower case characters without accents.

Note that the lemma may contain extra characters, for example:

Lemma	Occurs in	Meaning
"βάτος (I)"	Luke 6 : 44	Thorn bush or bramble
"βάτος (II)"	Luke 16 : 6	"Bath," a liquid measure

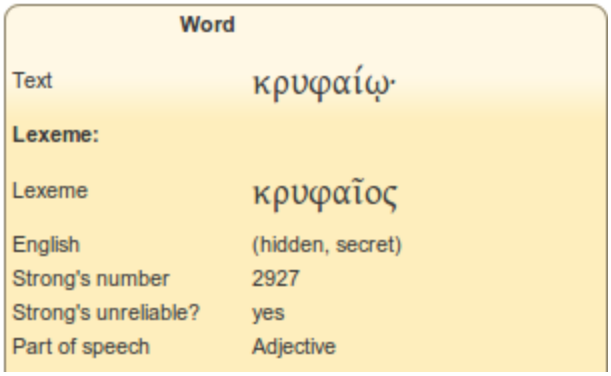
In the *raw_lemma* feature, both of these are given as "βατος".

C.1.3 Features: strongs, strongs_unreliable, and english

The *strongs* feature is an integer feature containing Strong's number for the lemma. The *strongs_unreliable* feature is a Boolean feature which is *true* if the indicated Strong's number is considered unreliable.

The *english* feature gives the English gloss for the lemma. This is based on Strong's number and taken from Jeffrey Dodson's Greek dictionary.

If a Strong's number is considered unreliable, Bible OL lists the corresponding gloss in parentheses. For example (from Matthew 6 : 18):



C.1.4 Feature: psp

The *psp* feature indicates the part of speech of the word. It is an enumeration features of type *psp_t* and can have these values:

- adjective
- adverb
- aramaic
- article
- cond
- conjunction
- correlative_or_interrogative_pronoun
- correlative_pronoun
- demonstrative_pronoun
- hebrew
- indefinite_pronoun
- interjection
- interrogative_pronoun
- letter_indeclinable
- noun
- noun_other_type_indeclinable
- numeral_indeclinable
- particle
- personal_pronoun
- possessive_pronoun
- preposition
- proper_noun_indeclinable
- reciprocal_pronoun
- reflexive_pronoun
- relative_pronoun
- verb
- NA (i.e., not applicable)

C.1.5 Features: person, number, gender, case, and possessor_number

The *person, number, gender,* and *case* features indicate the person, number, gender, and case, respectively, of the word. They are enumeration features of type *person_t, number_t, gender_t,* and *case_t* respectively.

For possessive pronouns the *number* feature indicates the number of the owned item, and the *possessor_number* feature indicates the number of the owning item. The *possessor_number* feature is an enumeration feature of type *number_t*.

The *person_t* enumeration has these values:

- first_person
- second_person
- third_person
- NA (i.e., not applicable)

The *number_t* enumeration has these values:

- singular
- plural
- NA (i.e., not applicable)

The *gender_t* enumeration has these values:

- masculine
- feminine
- neuter
- NA (i.e., not applicable)

The *case_t* enumeration has these values:

- nominative
- vocative
- genitive
- dative
- accusative
- NA (i.e., not applicable)

C.1.6 Features: tense, voice, and mood

The *tense, voice,* and *mood* features indicate the tense, voice, and mode, respectively, of a verb. They are enumeration features of type *tense_t, voice_t,* and *case_t* respectively.

The *tense_t* enumeration has these values:

- present
- imperfect
- future
- second_future
- aorist
- second_aorist
- perfect
- second_perfect
- pluperfect
- second_pluperfect
- NA (i.e., not applicable)

The *tense_t* enumeration has these values:

- active

- middle
- passive
- middle_or_passive
- middle_deponent
- passive_deponent
- middle_or_passive_deponent
- impersonal_active
- NA (i.e., not applicable)

The *mood_t* enumeration has these values:

- indicative
- subjunctive
- optative
- imperative
- infinitive
- participle
- imperative_participle

C.1.7 Feature: suffix

The *suffix* feature indicates the meaning of a word suffix. It is an enumeration features of type *suffix_t* and can have these values:

- superlative
- comparative
- interrogative
- negative
- attic
- particle_attached
- crasis
- NA (i.e., not applicable)

C.1.8 Feature: ref

The *ref* feature indicates the Bible verse to which this word belongs. For example, all words in Luke 2 : 5 have the *ref* feature set to "Luke 2:5".

C.1.9 Features: form_tag and functional_tag

The *form_tag* and *functional_tag* are taken directly from the original CSV file on which this Emdros database is built. They are not used by Bible OL.

C.2 The *sentence* Object

The *sentence* object has no features. Its purpose is merely to group words.

C.3 The *clause1* and *clause2* Objects

The *clause1* and *clause2* objects have a single feature, *typ.*

Although Bible OL only provides three levels of the syntax trees (*sentence, clause1,* and *clause2*), the Syntax trees on which the database is based contain considerably more levels. Adding more levels to Bible OL would require a completely different way to present the syntax trees, and three levels are considered adequate for most cases.

C.3.1 Feature: typ

The *typ* feature indicates the function of the clause. It is an enumeration features of type *clause_type_t* and can have these values:

Value	Meaning
ADV	Adverbial function
CL	Clause
IO	Indirect object function
O	Object function
O2	Second object function
P	Predicate function
S	Subject function
V	Verbal function
VC	Verbal copula function

C.4 The Other Object Types

For information about the other object types in the nestle4 database, please consult the MQL code used for generating the database. The MQL code is found in the file TBD.

C.5 The Origin of the Data

The current Emdros database comes from three sources:

- A CSV file containing the text and grammar information, provided to me by Ulrik Sandborg-Petersen.
- A lexicon derived from Jeff Dodson's Public Domain lexicon of the Greek NT, provided to me by Ulrik Sandborg-Petersen.
- Syntax trees downloaded from https://github.com/biblicalhumanities/greek-new-testament.

The text and the lexicon are also available from https://github.com/biblicalhumanities.

The Emdros database has been created based on these sources using my program *nestle2mql*, which is currently not published or documented.

Index