

# Generating nestle1904

Claus Tøndering

17 May 2023

## 1 Introduction

This document gives a brief introduction to how to generate the nestle1904 Emdros database for Bible Online Learner on a Linux computer.

## 2 Prerequisites

In order to execute this process, you need to *recursively* clone the GitHub repository <https://github.com/EzerIT/nestle1904.git>:

```
git clone --recursive https://github.com/EzerIT/nestle1904
```

In the following description I assume that you have cloned the repository into a folder called nestle1904.

## 3 Generating nestle1904

Executing the command “make” in the nestle1904/src folder should generate the nestle1904 Emdros database for Bible OL. This takes about two minutes on my computer.

The steps of the process are detailed in the following sections.

### 3.1 Compiling *nestle2mql*

The C++ source code for *nestle2mql* is compiled. A C++ compiler that supports the “-std=c++20” flag is required.

### 3.2 Executing *nestle2mql*

**Input:** ../nestle1904-1.2/nestle1904.csv

**Output:** nestle.mql

The file ../nestle1904-1.2/nestle1904.csv contains the text to the entire New Testament in tab-separated lines. The *nestle2mql* uses this to generate MQL code for the *word* objects of the database.

### 3.3 Compiling *add\_sentences/find\_sentences* and *add\_sentences/maketext*

The C++ source code for *find\_sentences* and *maketext* in the *add\_sentences* folder is compiled.

### 3.4 Executing *add\_sentences/maketext*

**Input:** ../greek-new-testament/syntax-trees/nestle1904/xml

**Output:** add\_sentences/xmlWithNode.txt

The *maketext* program reads XML files containing the sentence structure of the Greek NT and generates a file containing *nodeId:word* pairs from those files.

During the process tonos accent marks are replaced by oxia accent marks in the text. The reason for this is historical. Earlier versions of the XML files used oxia accent marks, and Bible OL still uses these marks.

For a discussion of the difference between tonos and oxia and a history of their use, see the section “A Note on Greek Accents in Unicode” in the chapter “Emdros Databases in Bible OL” in the Bible OL technical documentation.

### 3.5 Executing *add\_sentences/find\_sentences*

**Inputs:**

- ../greek-new-testament/syntax-trees/nestle1904/xml
- add\_sentences/xmlWithNode.txt

**Output:** add\_sentences/add\_sentences.mql

The *find\_sentences* program reads XML files containing the sentence structure of the Greek NT and generates an MQL file for the *sentence*, *clause1*, and *clause2* objects of the database.

### 3.6 Applying the MQL Files

**Inputs:**

- nestle1904.mql
- add\_sentences/add\_sentences.mql

**Output:** The nestle1904 Emdros database

The “mql” command is executed and creates the nestle1904 Emdros database from the MQL files generated in the previous steps.

### 3.7 Compiling *o2t* and *t2o*

The C++ source code for the programs *o2t* and *t2o* is compiled.

These two utility programs perform oxia-to-tonos and tonos-to-oxia conversion on their input files. They are not used in the generation of the nestle1904 database, but they may be useful in other contexts. (See Section 3.4.)

### 3.8 Compiling *hintsdb*

The C++ source code for *hintsdb* is compiled.

### 3.9 Executing *hintsdb*

**Inputs:**

- The nestle1904 database
- GREEK\_BibleOL\_nominal-ambiguity-project\_v1.21.csv
- AmbiguousVerbalForms20221021\_BibleOL-export.csv

**Output:** The `nestle1904_hints.db` database.

The *hintsdb* program is executed. It creates the so-called “hints database”. For information about this database, see the chapter *Hints* in the Bible OL technical documentation.

The `.csv` input files are taken from the corresponding `.xlsx` files created by Oliver Glanz. They contain information about various possible interpretations of a particular verbal or nominal form.

## 4 Generating Lexicons

There are currently no programs available for generating Greek lexicons.