# Test Document for XML Conversion

The following document is based on normal.dot; in all formats, Arial Unicode is used. A few paragraph styles have been added to demonstrate transfer of paragraph style names.

To keep the document clearly arranged the possible contents are enlisted with short comments. Generally, certain features of the text will not be evaluated. Most of all, this applies to the font because of the fact that a reliable conversion of the text solely is possible if only one single font is used containing as much as possible characters. Therefore, all paragraph styles are held in Arial Unicode. The actual font size equally will not be evaluated since the layout will result from the data structure later on.

# Headline Levels

The following chapter contains all convertible headline levels.

## Headline Level 2

The following chapters contain the remaining headline levels.

### Headline Level 3

The following chapters contain the remaining headline levels.

#### Headline Level 4

The following chapters contain the remaining headline levels.

##### Headline Level 5

The following chapters contain the remaining headline levels.

###### Headline Level 6

The following chapters contain the remaining headline levels.

Headline Level 7

The following chapters contain the remaining headline levels.

Headline Level 8

The following chapters contain the remaining headline levels.

Headline Level 9

The following chapters contain the remaining headline levels.

# Tables

In tables structures (rows, cells, and spanning of cells) and background color will primarily be transferred. Lines will not be transferred.

**Table 1:** Demographic Development in Göttingen 1740-1795

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Residents (without students)** | **Students** | **Professionals** |
| 1740 | 5562 | 313 | 670 |
| 1756 | 6900 | 600 | 698 |
| 1766 | 6306 | 695 | 702 |
| 1775 | 8107 | 840 | 804 |
| 1785 | 8674 | 838 | 826 |
| 1795 | 9078 | 674 | 833 |

If required background color of cells may be used to mark table header and footer. This has to be set up individually.

# Inline Elements

The following chapter covers all inline elements transferred actually. Theoretically, this may discretionary be extended.

## Character Formats

Following chapters contain all character formats considered during conversion. Other character formats are not prohibited but will be ignored during conversion.

Sample with text in **bold**, *italic*, ***bolditalic*** and underlined. Line designs will not be transferred in underlines. Superscript and subscript: 17 mg CO2 / m3.

Special distinctions Smallcaps, text with background color, strikethrough ~~text~~ and colored text. In texts with background color the information of ‘background color’ will be transferred, corresponding color will not be transferred.

Separately defined character formats like special will also be transferred. In this case not obvious situations may occur if character formats merely contain features that may also be allocated by means of a word function.

## Formulae

Integrated formula editor can compile Mathematical formulae from Word 2007 and higher and transfer as MathML[[1]](#footnote-1). Formulae compiled with former version of Word Formula editor cannot be transferred.

It is recommendable to use a separate paragraph style with auto line height when using out-of-line formulae.

It is possible to use inline formulae: unaesthetic better . In this connection, solely simple formulae expressions should be used; otherwise readability would suffer.

## Footnotes

Footnotes generated by standard range of functions of Word[[2]](#footnote-2) can be transferred in XML. Endnotes are not supported actually. However, there is no problem to add this functionality.

## Cross reference

Word supports several variations of cross references. The most important base on automatically numbered elements. Since headlines are not numbered in this document this functionality does not exist. Therefore, a reference only can contain the headline text. These are references with headline text on Headline Levels and Tables. In principle, pagination may be indicated like this 2 but it is not that practicable.

Beyond this, bookmarks may be used. You have to add a bookmark at the target that you want to reference. Here are references to headline levels and tables.

## Index Entries

You may add index entries by Word function for index entries. Additional functions "see", bold pagination and italic pagination will be transferred.

# Lists

The following chapter contains the most important list types. **Note:** Numbering values of lists will not be transferred i.e. a new list always starts with 1 irrespective of what Word shows. Numbering of lists from Word is always good for a surprise. It is possible that conversion of lists does not always lead to the favored target. This is not a converter’s problem.

From here, uncounted lists:

* First list point level 1
* Second list point level 1
  + First list point level 2
  + Second list point level 2
    - First list point level 3
    - Second list point level 3
    - Third list point level 3
  + Third list point level 2
* Third list point level 1
* Forth list point level 1

From here, counted lists:

1. First list point level 1
2. Second list point level 1
   1. First list point level 2
   2. Second list point level 2
      1. First list point level 3
      2. Second list point level 3
      3. Third list point level 3
   3. Third list point level 2
   4. Forth list point level 2
3. Third list point level 1
4. Forth list point level 1

Actually lists are supported up to max. 3 level.

# Images

Theoretically, images may be transferred from a Word document. For print products the quality of word internal Jpeg data is generally not accepted. For this reason, this version of the converter does not support it. Basically, read-out of word internal image data is possible.

The most common possibility to transfer images in XML is to mark data name and image underline by a special paragraph format and to implement this within the individual conversion accordingly.

1. This format is the standard for XML but receives limited support by the most important layout programs. [↑](#footnote-ref-1)
2. You will find this under reference -> insert footnote. [↑](#footnote-ref-2)