

WordTeX

A WYSIPCTWOTCG¹ Typesetting Tool

Tom Wildenhain

Abstract

`WordTeX` is a plugin for Microsoft Word that attempts the impossible: creating documents that appear to be written in `LATEX` while irritating people who like `LATEX`. It is both stupidly impractical and surprisingly useful, offering an editing experience that is initially more enjoyable than `LATEX` and Word but is asymptotically more complicated than either. In this paper, I will explain the results of my `WordTeX` research. I will occasionally include content that has no relevance to the paper and is simply used to showcase how `WordTeX` renders certain elements. $(x + 1)^2 = x^2 + 2x + 1$.

1 Introduction

`LATEX` is a popular typesetting tool for creating complicated, consistently-formatted documents. Many students and scientists use `LATEX` because the finished documents have a clean, professional look that is hard to achieve in standard word processors.^[citation needed] However, these high-quality results come at the cost of a steep learning curve, potentially tedious editing experience, and sudden anxiety when homework assignments won't compile before a deadline.

Microsoft Word is a WYSIWYG editor, enabling users to have confidence in the appearance of their documents throughout the editing process. In fact, (real) research shows that Word novices are more productive and

¹ What You See Is Pretty Close To What Other Tools Can Get

make fewer errors than L^AT_EX experts when creating certain types of documents.²

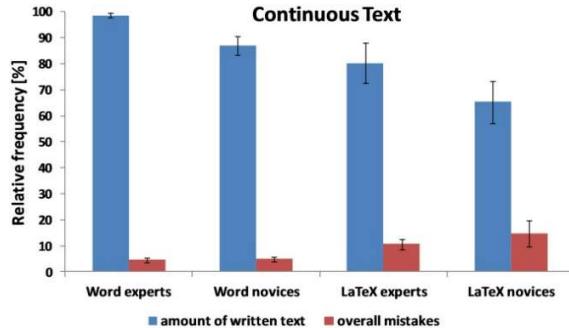


Figure 1: Word users are more productive than L^AT_EX users for some document types. (Knauff & Nejasmic 2014)

However, when authoring more complicated documents, Word users spend much of their time clicking through menus and fiddling with formatting, leading to inconsistent document structure.

WordT_EX aims to combine the strengths of Word and L^AT_EX, creating high-quality documents in real time, while annoying both Word and L^AT_EX fans. It is particularly useful for typesetting homework assignments, although students should be warned that professors may warn students that “using L^AT_EX is an important skill that will help you in the long term.”

2 The WordT_EX Template

The main component of WordT_EX is the WordTeX.dot template file. The template includes a set of styles that closely approximate the appearance of L^AT_EX. The template is designed for Microsoft Office Word 2016 for Windows (or Office 365 ProPlus) and might not work correctly in other versions of Word.³

² Knauff, M., & Nejasmic, J. (2014). An Efficiency Comparison of Document Preparation Systems Used in Academic Research and Development.

³ I am not yet sponsored by Microsoft.

2.1 Fonts

Word_{TEX} users should install the LM Roman fonts and Latin Modern Math font (published by GUST) before using the template.

2.2 Styles

Word_{TEX} uses styles to maintain a consistent look throughout the document. Styles are provided for the title, subtitle, inline code, headings, and other common L^AT_EX elements (Figure 2).



Figure 2: The style gallery

A user can apply a style by selecting it from the style bar or by typing **ctrl+shift+S** and then the name of the style (the latter is much faster). Typing **ctrl+shift+Z** clears character styles (which are applied within a line). Switching to the Normal style clears paragraph styles. Heading styles are automatically numbered, and styles throughout the document are updated if a style is modified.

2.3 Math

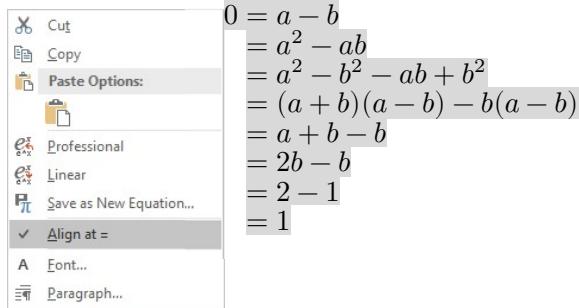
Math in Word_{TEX} uses the Latin Modern Math font. To enter or exit math mode a user can type **alt+=**. Latex commands like `\sum` and `\subseteqq` are supported and render in real time. Word uses parentheses for grouping instead of curly braces. For example, `\sum_{i=1}^{10}` `i=55` renders as:

$$\sum_{i=1}^{10} i = 55$$

2.3.1 Alignment

A set of aligned equations can be created by selecting “align at =” from the equation context menu.

Let $a = b \neq 0$.



2.3.2 Functions

Functions like $\sin(x)$ and $\log_2(x)$ are automatically written without italics if they appear in the recognized function list. The equation options dialog (accessible from the conversions section) can be used to add more functions to the list.

2.3.3 Blackboard Bold

Blackboard bold letters like \mathbb{R} and \mathbb{N} can be inserted by typing `\doubleR` or `\doubleN`. They unfortunately look different from the L^AT_EX Blackboard Bold font, since they use glyphs from the Latin Modern Math font. Future researchers might be able to edit the font to include the `\mathbb` glyphs. The commands for these symbols can be shortened (see the section on macros).

2.4 Proofs

Proofs are started with the word “proof” set in the proof character style. To end a proof, a user can type `tab` and then `\qed` to insert a \square symbol.

Proof. This is a proof. It is written in WordT_EX. \therefore WordT_EX can make proofs. \square

2.5 Code

```
def print_code_instructions():
    assert includes_styles(inline and block_code)
    # Syntax highlighting support is planned and will
    # hopefully be released soon.
```

2.6 Macros

\LaTeX includes a powerful macro system allows users to define custom commands. Word’s AutoCorrect is somewhat similar, but only performs basic text replacement. For example, you can add a math AutoCorrect entry that replaces \R with \mathbb{R} , but AutoCorrect entries can’t take arguments.

Despite these limitations, AutoCorrect is in fact exactly as powerful as \LaTeX macros, as they are both Turing Complete.⁴ We can easily simulate a Turing Machine using AutoCorrect entries by representing the state as a string that “reads” characters by adding different entries for every combination of adjacent characters.

\$ \stateA 1 0 1 1 0 0 1 1 0 1 \$	
\$ \stateB 0 1 1 0 0 1 1 0 1 \$	
\$ 0 1 1 0 0 1 1 \stateB 0 1 \$	
\$ 0 1 1 0 0 1 1 0 \stateC \$	
\$ 0 1 1 0 0 \stateC 1 1 0 \$	

\stateA 1	\stateB
\stateB \$	\stateC \$
\stateB 0	0 \stateB
\stateB 1	1 \stateB

Figure 3: An example computation trace. \$ marks the ends of the input.

Unfortunately, AutoCorrect macros only evaluate once from left to right, so the Turing Machine will stop running if it ever moves to the left. To continue evaluation, a user can repeatedly press the Convert button in the Equation Tools tab. Be warned: if the Turing Machine moves to the right without halting, evaluation will not stop, and Word will freeze. I do not know if Microsoft is aware of this issue.

2.7 Printing

While Word has a built-in export to PDF option, it unfortunately does not embed otf fonts (like the Latin Modern fonts). The best option is to print to the Microsoft Print to PDF printer, which will embed the fonts.

⁴ <https://www.sharelatex.com/blog/2012/04/24/latex-is-more-powerful-than-you-think.html>

3 Conversions

While **WordTeX** is superior to **LATEX** in many ways, sometimes **LATEX** source is required for a conference or assignment.⁵ Thanks to Pandoc, **WordTeX** files can be converted to **LATEX** source code.⁶ Mathematical expressions ($\int_0^4 x dx = 8$), **inline code**, and most **formatting** is converted. In fact, here's the source for this paragraph:

```
\hypertarget{conversions}{%
\section{Conversions}\label{conversions}}}

While WordTeX is superior to LaTeX in many ways, sometimes LaTeX source
is required for a conference or assignment.\footnote{Fortunately,
SIGBOVIK does not have such ridiculous restrictions.} Thanks to
Pandoc, WordTeX files can be converted to LaTeX source code.
\footnote{\url{https://pandoc.org/}}
Mathematical expressions (\(\int_0^4 x dx = 8\)),
\texttt{inline\ code}, and most \emph{formatting} is converted.
In fact, here's the source for this paragraph:
```

The **WordTeX** plugin adds **Copy as LaTeX** and **Paste From LaTeX** buttons to Word. They use Pandoc to convert between formats in real time. The conversion isn't perfect, but is fairly close.

4 Similarity to **LATEX**

4.1 Experiment

I conducted a double-blind randomized study to determine whether documents typeset using **LATEX** and **WordTeX** are distinguishable (Fig. 4). The test subject was blindfolded, and the experimenter (also blindfolded) told her to write L or W on each paper she believed to be a **LATEX** or **WordTeX** document, respectively. Data was collected until the results supported the hypothesis that **LATEX** and **WordTeX** are indistinguishable.

⁵ Fortunately, SIGBOVIK does not have such ridiculous restrictions.

⁶ <https://pandoc.org/>



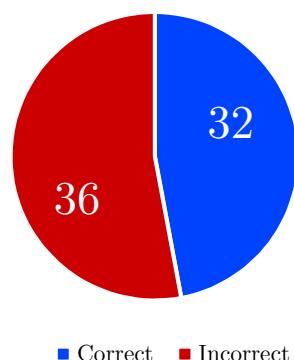
Figure 4: Experimental setup

4.2 Results

The papers were misidentified most of the time.

	L ^A T _E X	WordT _E X
Identified as L ^A T _E X	16	18
Identified as WordT _E X	18	16

It is clear from the following chart that the red side is bigger than the blue side ($p < 0.5$).



4.3 Conclusions

Running a χ^2 test, the p -value is 0.628, which is not significant. Therefore, documents typeset in L^AT_EX do not appear to be significantly different from those typeset in WordT_EX.

5 Summary

WordT_EX is a typesetting system that supports the basic functionality of L^AT_EX while utilizing the editing convenience of Word. Word's Turing-complete macros ensure that WordT_EX is just as powerful as L^AT_EX (for questionable definitions of "powerful"). The WordT_EX plugin allows for easy conversion between WordT_EX and L^AT_EX. Experimental results suggest that WordT_EX and L^AT_EX documents are indistinguishable. In light of these results, I encourage all scientists, students, and professors to abandon L^AT_EX immediately and use WordT_EX for future work.