

Re: English versus German

Source: <http://sci.tech-archive.net/Archive/sci.lang/2009-07/msg00204.html>

- *From:* "Peter T. Daniels" <grammatim@xxxxxxxxxxxx>
 - *Date:* Fri, 3 Jul 2009 05:41:32 -0700 (PDT)
-

On Jul 2, 11:50 pm, Nathan Sanders <nsand...@xxxxxxxxxxxx> wrote:

In article

<b90de218-898c-4f41-8e2c-99d8c8065...@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>, "Peter T. Daniels" <gramma...@xxxxxxxxxxxx> wrote:

On Jul 2, 8:15 pm, Nathan Sanders <nsand...@xxxxxxxxxxxx> wrote:

In article

<3598a0f9-7f6d-43ac-9e3b-12d991b6b...@xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx>, "Peter T. Daniels" <gramma...@xxxxxxxxxxxx> wrote:

On Jul 2, 5:00 pm, Nathan Sanders <nsand...@xxxxxxxxxxxx> wrote:

No ligatures for fi, fl, ff, ffi,
or ffl.

A point gone unrebutted.

Covered by "not the writer's business."

Kerning is terrible. "Table"
has the wrong spacing
between T and a,
and tall italics characters
inside parentheses bump
into the
right-hand parenthesis.

Another point gone unrebutted.

Re: English versus German

Covered by "not the writer's business."

Multiple
accent
marks
on
the
same
character
(e.g.,
acute
over
macron).

Possible in
any system
that
supports
Unicode.

TeX was able to handle this
easily before Unicode.

Uh, so what? This is 2009.

The point is that the way TeX does it doesn't require
Unicode, doesn't
require the Combining Diacritical Marks range, doesn't
require
assigning keyboard shortcuts, or any of that other fancy stuff.
It
works and has always worked directly "out of box", in a
simple
straightforward manner.

Except that it's utterly incompatible with any user who doesn't happen
to have TeX.

You clearly don't understand what "plain text" means.

Re: English versus German

It means 'text that is not encrypted'. The phrase "plain text" doesn't appear in what you wrote above. If you meant "text file," i.e. .txt, that's text using only the ASCII symbols (codepoints 000–127).

Yes, there are markup commands in the file, but the vast majority of TeX source is just plain ordinary legible text.

A text in Unicode can be read in any current word processor (except yours).

I have the latest version of Word.

Your TeX, idiot.

It still handles
it easily: `\|=a` puts an acute
over a macron over "a". I
know how to
get an acute accent over "a"
in Word, and I've modified
my keyboard to
allow me to get a macron
(this is not an
out-of-the-box ability!),
but
I don't know of a simple
way to get both accents on
the same character
in Word.

Have you not discovered the Insert Symbol
button?

What's the keyboard shortcut for that?

Because it can be added to the QAT (Quick Access Toolbar), you can assign any shortcut you want to it. I've seen no reason to do so because it's a GUI, so the mouse is involved anyway.

Re: English versus German

Ah, yes, taking the hands off the keyboard is a highly efficient way to work!

Only for rare symbols that I don't see the need to assign keyboard shortcuts to. I can of course keyboard-type anything I need with a macron, a breve, a hachek, an underdot, an overdot -- basically, anything encountered in Semitic or Indo-European transliterations, and pinyin; if I need to type stuff using different transliteration conventions, I can easily add shortcuts for those characters as well. (For Vietnamese, obviously, I switch to the Vietnamese IME.)

Presumably you could also assign a keyboard shortcut to it in Word2003 as well, but I'm not about to wake 2003 up to check.

Of course not. Waking up any version of Word is a chore!

No; I haven't bothered doing the registry change that allows two versions to operate on the same computer side by side without having to "reconfigure" when the one that wasn't last used is opened. But if your Word doesn't open nearly instantaneously when you click it in your Start menu [I suppose you have a "Dock" instead] or double-click a .doc(x) file, then your computer is broken.

And if you haven't even discovered the Combining Diacritical Marks range of Unicode, you have very, very little business criticizing its linguistic abilities.

Oh, I know where it is. Unfortunately, they don't all combine properly in Word. I just tried "a" + combining macron + combining acute and got a misaligned macron and a misaligned acute superimposed on each other (rather than stacked and properly centered, as occurs in TeX).

That's not Word's fault, it's the font's fault. Times New Roman has never failed me in adding a diacritic to a letter (it even knows to

Re: English versus German

raise them above caps and tall letters.)

I tried it with Times New Roman, too. Word failed even worse! The combining macron shows up as a square block (indicating a missing symbol). The combining acute looks fine on "a", but looks terrible on "m": it's too high, and off-center.

Then you need to update your font. Or maybe your OS isn't compatible with the latest version of Unicode (that's why I had to switch to Vista).

I just put an acute accent over an "m" in Word, and it changed the font from Cambria to MS Reference Sans Serif, and now, everything I type after that is in the new font!

Then you should learn how to use Word properly.

Praytell, how do you prevent the font from switching when the accented character try to type doesn't exist in the font? How do you know, a priori, whether any particular accent+character combination is missing?

You could look in Character Map. (Or, of course, in Insert Symbol.) That shortcoming has nothing to do with Word.

Of course it does. If the combined character doesn't exist in the font, but is easily created (e.g., by taking an acute accent and placing it in the correct position over an "m"), then the proper behavior is to create the character, not change the font for the rest of the document!

No one using Word2007 in Vista experiences such a problem.

Re: English versus German

If you don't like Times New Roman, you could try the Lucida family, which was designed by Chuck Bigelow in consultation with Bill Bright.

Lucida Bright was the worst of all! Both of the combining diacritics showed up as "missing character" squares.

Not "Lucida Bright." "Lucida Unicode" or "Lucida Sans" and I think a couple other varieties.

Lucida Grande puts the acute on "a" correctly, and at least gets the right height for the acute on "m", but it's still off-center. And when trying to combine macron with acute, I still get superposition, not stacking.

Maybe "Grande" was a compromise to get around shortcomings in the Mac OS. I haven't heard of it. (I also don't know anything about Lucida Bright. Maybe it derives from a PostScript font made for American Indian and/or Dravidian work.)

When TeX can't find the combination, it creates it from scratch, and the result looks like it belongs with the rest of the text.
When Word can't find it, it changes the font.

Considering Knuth's talent for typography, I think we can consider that yet another win for Word. It's its way of telling you you need a better-supplied font. But it only does that for precomposed characters (i.e. with Unicode assignment) that are missing from the font in question.

I don't see how on earth you can think this is a win for Word!

Because I have no problem with diacritized characters (and, obviously, I deal with the whole world's) and you do.

`\m` in TeX produces "m" (in the same font as everything else) with an acute accent at the correct height, perfectly centered over the "m".

Re: English versus German

Neither "option+e m" or "m + combining acute" work properly in Word for any of the expected fonts (default, Times New Roman, various Lucidas).

"option-e m" is Mac-specific. Windows doesn't have system-wide shortcut standards. In Word, the acute shortcut is Ctrl-' (apostrophe) with the French vowels.

I just tested it -- m + combining-acute gives a proper character in Times New Roman, and AFAICT there's no precomposed glyph for it in Unicode, so it seems the typography routines in the Mac OS are faulty. It also positions it properly over an M.

TeX's behavior is far more desirable.

Not to people who care what their output looks like.

I definitely care that my acute accents are properly centered!

Then maybe you should switch to Windows Vista.

If the accented character is not a pre-composed character in the font (as acute-m is not in Cambria), TeX will compose the accented character on the fly. There is no font change for either that character or the following ones.

Just like Word!

Nonsense. I just say the font-changing behavior in Word.

Re: English versus German

If m–acute is a Unicode character, then it's telling you you can't have that character in that font.

A shame. Apparently with Word, I can't have that character in any font!

You can have it in Times New Roman. I'll check a few other fonts ...

Works: Arial, Adobe CJK fonts [they came with CS4], Aharoni, Helvetica, Tahoma, some (but by no means all) of the romans that are part of exotics (such as Hebrew and Arabic fonts); all CJK fonts.

Doesn't work: Calibri, Cambria, display fonts generally, Lucida (all the varieties I have).

It is thus up to the type designer whether to include the capability. It also doesn't relate to OpenType vs. TrueType.

Fortunately, TeX let's me have it with any font I want!

Did you test it with, say, Copperplate Gothic?

Can you send a Word file to a colleague who doesn't own the same version of Word as you (or no version at all!), and/or is working on a different OS than you (Mac versus PC... god forbid he's on Linux!), and he'll be able to open your file and it will look the same as on your machine? As recently as last year, that wasn't true for me when I was collaborating with a colleague. Random symbols were missing, pagination was different, and even the fonts weren't consistent.

Re: English versus German

If the same fonts were installed on both computers, the fonts were not a problem. If the document was assigned to the same printer driver on the different computers, there would be no formatting problem.

We had the same fonts. The "Insert symbol" function is/was broken.

If the inserted symbol was done on a PC, it showed up as a box on a Mac.

You failed to mention that you were using obsolete, pre-Unicode fonts.

I was using the standard fonts that come with the latest version of Word, and the font in question was Times New Roman.

Macs and PCs differ in the lowest range of non-ASCII characters.

None of these are problems with TeX, because TeX source files are all plain text files, the most basic text file you can possibly have, free and common to every computer system, every operating system, and every time period.

And therefore consuming lots of extra computing power every time they're opened.

?!?! Plain text files are the easiest files for a computer to open!

Re: English versus German

But they will be gibberish anywhere else, since they have all those thousands of coding characters (sort of like WordPerfect in its pre-WYSIWYG days, but with much more code).

Apparently you don't know what "plain text" means. Here are some sample paragraphs from my handout from this year's LSA, written in LaTeX:

And you had to type all those extra codes? What a waste.

read more »-

Maybe.

.