State of CJK issues of LibreOffice, 2020 edition

Shinji Enoki
shinji.enoki@libreoffice.org

Shinji Enoki
shinji.enoki@libreoffice.org

OpenSUSE + LibreOffice
Conference ONLINE
15 Oct. 2020

Agenda

- What are CJK issues?
- Typical LibreOffice CJK functions
- LibreOffice CJK bug status
- Typical CJK issues
- Conclusion

Shinji Enoki（榎真治）

- Member of LibreOffice Japanese Team(2011-)

- Member of The Document Foundation (2014-)

- Activity: organizing events, grows community, sometimes QA

- Other community:
  - Vice-chairperson of Japan UNIX Society
  - Staff of KANSAI OPEN SOURCE
  - Nextcloud, OpenData, etc…

Live in Nishinomiya, Japan
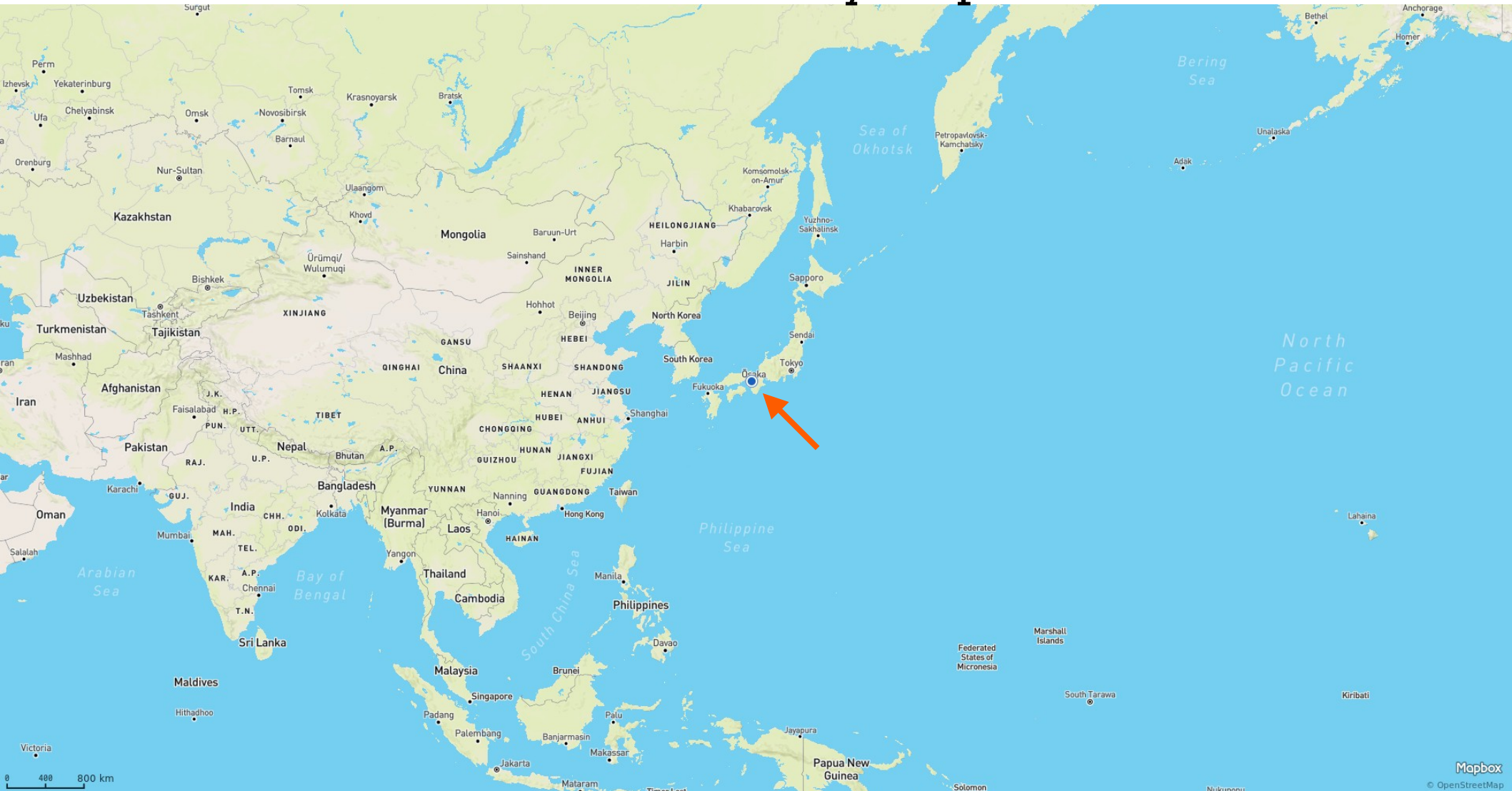
What are CJK issues?

What are CJK issues / bugs

- CJK is an abbreviation for "Chinese-Japanese-Korean"

- Chinese, Japanese, and Korean are different languages, but they have some common features

- LibreOffice has many language-specific features and issues, CJK issue is one of them

- CJK issues happen when used in CJK's environment

TDF : Next Decade Manifesto

"To support the preservation of mother tongue by encouraging people to translate,document, support, and promote our office productivity tools in their mother tongue"

• I think these "Our Values" is very important

CJK users need to report / reproduce CJK bugs

- It is hard to notice / understand CJK bug unless it is people of CJK

- Many LibreOffice developers are not CJK people

Typical LibreOffice CJK functions

Overview of CJK functions

- Text Layout
  - Vertical writing
  - Phonetic guides (ruby)
  - Line Composition
- Multibyte character
- Input methods
- ....
- CJK needs similar functionality, but its rules are different

Vertical writing (1)

- Supported among various functions
- When the function is changed, it sometimes breaks

Vertical writing (2)

- Writer: [Page Style] setting can be set to vertical writing

Page Style: Default Page Style

| Footer | Borders | Columns | Foot |
| Organizer | Page | Area | Transp |

Paper Format

Format: A4

Width: 21.00 cm − +

Height: 29.70 cm − +

Orientation: ○ Portrait
　　　　　　 ○ Landscape

Text direction: Left-to-right (horizontal) ▼   Paper tray: [Fro

Left-to-right (horizontal)
Right-to-left (vertical)
Left-to-right (vertical)

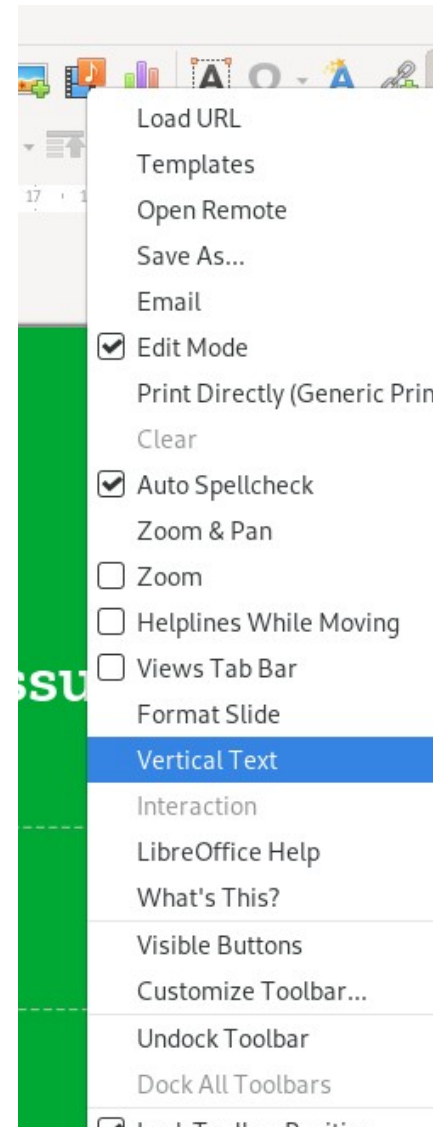**Margins**   **Layout Settings**

Left: Page layout: Right

ある日の暮方の事である。一人の下人が、羅生門の下に、雨やみを待っていた。

広い門の下には、この男のほかに誰もいない。ただ、所々丹塗の剥げた、大きな円柱に、蟋蟀が一匹とまっている。羅生門が、朱雀大路にある以上は、この男のほかにも、雨やみをする市女笠や揉烏帽子が、もう二三人はありそうなものである。それが、この男のほかには誰もいない。

何故かと云うと、この二三年、京都には、地震とか辻風とか火事とか饑饉とか云う災がつづいて起った。そこで洛中のさびれ方は一通りではない。旧記によると、仏像や仏具を打砕いて、その丹がついたり、金銀の箔がついたりした木を、路ばたにつみ重ねて、薪《たきぎ》の料《しろ》に売っていたと云う事である。洛中がその始末であるから、羅生門の修理などは、元より誰も捨てて顧る者がなかった。するとその荒れ果てたのをよい事にして、狐狸が棲む。盗人が棲む。とうとうしまいには、引取り手のない死人を、この門へ持って来て、棄てて行くと云う習慣さえ出来た。そこで、日の目が見えなくなると、誰でも気味を悪るがって、この門の近所へは足ぶみをしない事になってしまったのである。

Vertical writing (2)

- Vertical text box
  - Supported by Writer, Calc, Impress, Draw
  - Note: Different from normal text box
  - Right click on the toolbar-> [Vertical Text]

縦書きの
Vertical text box

Load URL
Templates
Open Remote
Save As...
Email
☑ Edit Mode
Print Directly (Generic Prin
Clear
☑ Auto Spellcheck
Zoom & Pan
☐ Zoom
☐ Helplines While Moving
☐ Views Tab Bar
Format Slide
Vertical Text
Interaction
LibreOffice Help
What's This?
Visible Buttons
Customize Toolbar...
Undock Toolbar
Dock All Toolbars

Phonetic guides (ruby)

- In Japanese and Chinese, mainly

  indicate how to read

Phonetic guides (ruby)

Writer: menu

[Format]>

[Asian Phonetic guide]

Line breaking rules(Line Composition)

- Paragraph style dialog ->Asian typography
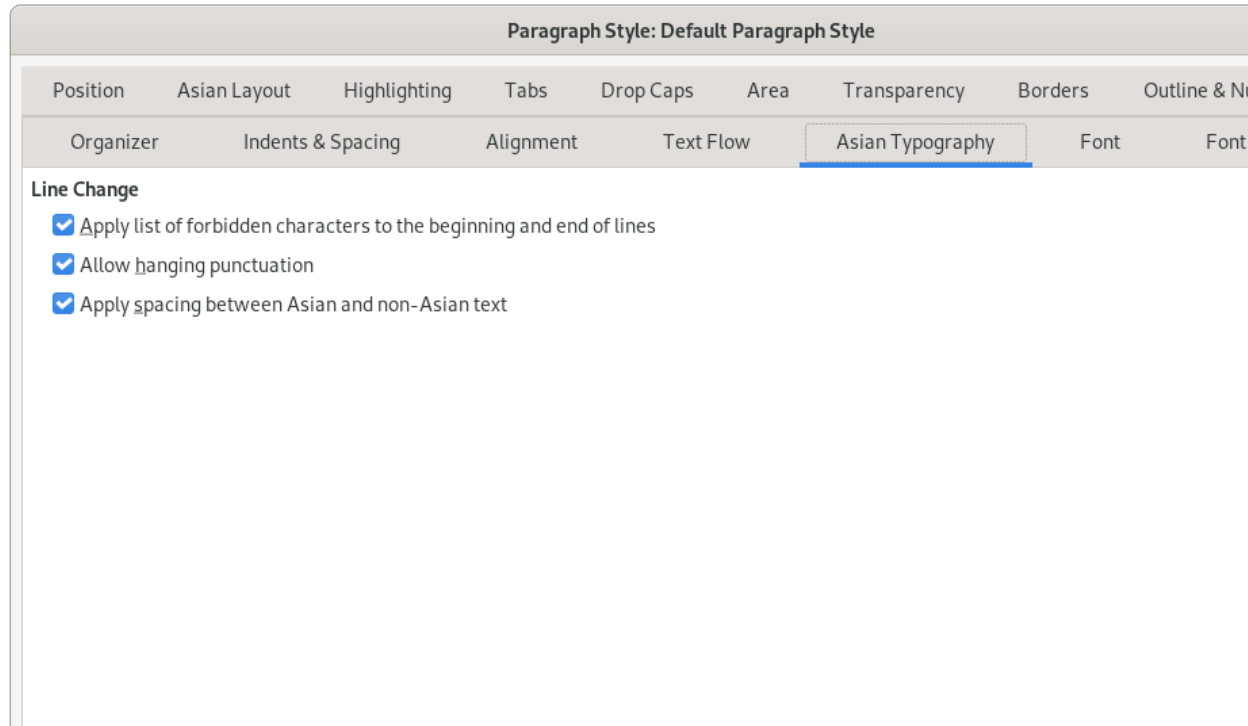
-

Paragraph Style: Default Paragraph Style

| Position | Asian Layout | Highlighting | Tabs | Drop Caps | Area | Transparency | Borders | Outline & Nu |

| Organizer | Indents & Spacing | Alignment | Text Flow | Asian Typography | Font | Font |

Line Change

- ☑ Apply list of forbidden characters to the beginning and end of lines
- ☑ Allow hanging punctuation
- ☑ Apply spacing between Asian and non-Asian text

　広い門の下には、この男のほかに誰もいない。ただ、所々｜丹塗《にぬり》の剥《は》げた、大きな円柱《まるばしら》に、蟋蟀《きりぎりす》が一匹とまっている。羅生門が、朱雀大路《すざくおおじ》にある以上は、この男のほかにも、雨やみをする市女笠《いちめがさ》や揉烏帽子《もみえぼし》が、もう二三人はありそうなものである。それが、この男のほかには誰もいない。

　広い門の下には、この男のほかに誰もいない。ただ、所々｜丹塗《にぬり》の剥《は》げた、大きな円柱《まるばしら》に、蟋蟀《きりぎりす》が一匹とまっている。羅生門が、朱雀大路《すざくおおじ》にある以上は、この男のほかにも、雨やみをする市女笠《いちめがさ》や揉烏帽子《もみえぼし》が、もう二三人はありそうなものである。それが、この男のほかには誰もいない。

　広い門の下には、この男のほかに誰もいない。ただ、所々｜丹塗《にぬり》の剥《は》げた、大きな円柱《まるばしら》に、蟋蟀《きりぎりす》が一匹とまっている。羅生門が、朱雀大路《すざくおおじ》にある以上は、この男のほかにも、雨やみをする市女笠《いちめがさ》や揉烏帽子《もみえぼし》が、もう二三人はありそうなものである。それが、この男のほかには誰もいない。

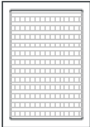# Text Grid

Page Style: Default Page Style ✕

Organizer	Page	Area	Transparency	Header
Footer	Borders	Columns	Footnote	Text Grid

Grid

○ No grid
○ Grid (lines only)
● Grid (lines and characters)

Grid Layout

Lines per page: 38 − + (1 - 38) Max. base text size: 11.00 pt − +

Characters per line: 43 − + (1 - 43) Max. Ruby text size: 8.00 pt − +

☐ Ruby text below/left from base text

Grid Display

☑ Display grid   Grid color: #C0C0C0 ▼
☑ Print grid

Help   Reset Apply Cancel OK

人の下人《げにん》が、羅生門《らしょうもん》の下で雨やみを

かに誰もいない。ただ、所々｜丹塗《にぬり》の剥《は》げた、

悉蟀《きりぎりす》が一匹とまっている。羅生門が、朱雀大路

この男のほかにも、雨やみをする市女笠《いちめがさ》や揉烏

三人はありそうなものである。それが、この男のほかには誰もい

京都には、地震とか辻風《つじかぜ》とか火事とか饑饉とか云

った。そこで洛中《らくちゅう》のさびれ方は一通りではない。

砕いて、その丹《に》がついたり、金銀の箔《はく》がついたり

薪《たきぎ》の料《しろ》に売っていたと云う事である。洛中

の修理などは、元より誰も捨てて顧る者がなかった。するとその

瓜狸《こり》が棲《す》む。盗人《ぬすびと》が棲む。とうとう

人を、この門へ持って来て、棄てて行くと云う習慣さえ出来た。

そこで、日の目が見えなくなると、誰でも気味を悪るがって、この門の近所へは足ぶみをしない

事になってしまったのである。

Reference: W3C Requirements documents

- Requirements for Japanese Text Layout
  - https://www.w3.org/TR/jlreq/
  - "This document describes requirements for general Japanese layout realized with technologies like CSS, SVG and XSL-FO. "
- Requirements for Chinese Text Layout 中文排版需求
  - https://www.w3.org/TR/clreq/
  - "This document was developed by people working in different areas, using both Simplified and Traditional Chinese. "
- Requirements for Hangul Text Layout and Typography
  - https://www.w3.org/TR/klreq/
- These three documents were updated in 2020

LibreOffice CJK bug status

Home | New | Browse | Search | [Search]   [?] | Reports | Help | New Account | Log In | Forgot Password

Bug 83066 (CJK) - [META] CJK (Chinese, Japanese, Korean, and Vietnamese) language issues

Status: NEW

Alias: CJK

Product: LibreOffice
Component: LibreOffice (show other bugs)
Version:
(earliest affected) unspecified
Hardware: All All

Importance: medium normal
Assignee: Not Assigned

URL:
Whiteboard:
Keywords:

Depends on: 35208 36709 39935 43671 50349 50901 55954 63011 65266 66772 66791 69809 71329 75790 77803 81272 81484 82115 82116 82293 82689 82690 82692 82930 83844 88752 89422 89979 90874 93695 94631 101679 101895 104378 104503 104927 105501 107195 107466 107467 107484 107485 108042 109030 113048 113184 113189 113191 CJK-Chinese-Traditional CJK-Chinese-Simplified CJK-Japanese CJK-Korean 114272 114520 114746 114760 114761 114763 114764 115321 115970 117554 118260 118476 118683 118849 119219 120171 126169 126406 127010 128080 128221 128286 129810 129940 131751 *132517* 132544 *134096* *134742* *136081* 136283 137393 40091 40335 44784 47719 49073 57139 58604 62750 64975 66126 67814 70097 72042 76255 76291 77076 77490 77514 77971 *77972* 78789 79766 80272 80346 80473 80513 80604 80788 80812 81144 82176 82818 83060 83145 83178 83252 83559 83752 83755 84138 85795 85954 87626 88169 88527 88855 89409 91288 92655 92782 94008 94066 95379 95656 95836 96091 96197 96457 96458 96459 96829 97405 97702 98165 98269 98527 98879 101729 102041 102074 102253 103528 103710 103718 103785 103850 104625 104726 104801 104828 104954 106097 106295 107056 107383 107435 107462 107487 109260 111717 111967 112066 112477 112593 112745 112807 112879 113481 114486 114736 115130 115353 116577 117857 118564 119211 120155 121161 121527 121855 122459 123956 123957 124657 125404 125446 125620 126157 127491 128276 128304 130077 130140 130193 130750 134566 136595

Blocks: Languages
Show dependency tree / graph

Reported: 2014-08-25 14:50 UTC by Matthew Francis

Modified: 2020-10-11 13:51 UTC (History)

CC List: 17 users (show)

See Also: Font-Rendering
Vertical-Text
RTL-CTL
Ruby
Text-Grid

Crash report or crash signature:

Depends on CJK Meta issue #83066

	2017/10	2018/9	2019/9	2020/9
Depends on: (All)	135	168	186	207
Open	59	73	73	80
Close	76	95	113	127

Meta issue for CJK

- Bug 83066 : [META] CJK (Chinese, Japanese, Korean, and Vietnamese) language issues

- Meta issue for each CJK language

- Bug 113193 : [META] Traditional Chinese (zh_TW, zh_HK)

- Bug 113194 : [META] Simplified Chinese (zh_CN)

- Bug 113195 : [META] Japanese CJK issues

- Bug 113196 : [META] Korean

Typical CJK issues

ODS exported by Excel with Asian phonetic guide is displayed as empty :Bug#129940

Ruby is not supported in Calc

Not only ruby, ruby base text

is not displayed

content.xml of ODS file

```
...e= uaca style:column-width= 1.964375cm /></style:s
s style:row-height="20pt" style:use-optimal-row-height
lly="table" style:master-page-name="mp1"><style:table-
automatic-styles><office:body><office:spreadsheet><tab
:o-whole-cell="true" table:use-wildcards="true" table:
able:name="Sheet1" table:style-name="ta1"><table:table
ame="ce1"/><table:table-row table:style-name="ro1"><ta
e-name="ru1"><text:ruby-base>小笠原</text:ruby-base><tex
.><table:table-cell office:value-type="string" table:s
ruby-text text:style-name="T1">ナルヒコ </text:ruby-tex
="16382"/></table:table-row><table:table-row table:num
peated="16384"/></table:table-row></table:table></offi
```

ruby include

ファイル(F)　編集(E)　表示(V)　挿入(I)　書式(O)　ス

	A	B	C	D
1				
2				
3				

ruby not include

ファイル(F)　編集(E)　表示(V)　挿入(I)　書式(O)　ス

A1　　　fx Σ ▾ = 小笠原

	A	B	C	D
1	小笠原			
2				
3				
4				

Not fixed, Xisco says Regression

When Skia is enabled in vertical writing, changing the font scale changes the font size: Bug#136081

Setting to change Font width,
Font size has been changed

Skia rendering is enabled
on Windows only

Not fixed

When Skiya is on, the 3-point leader(dot-dot-dot) is not working vertical writing

In vertical writing, the text does not fit in the text grid

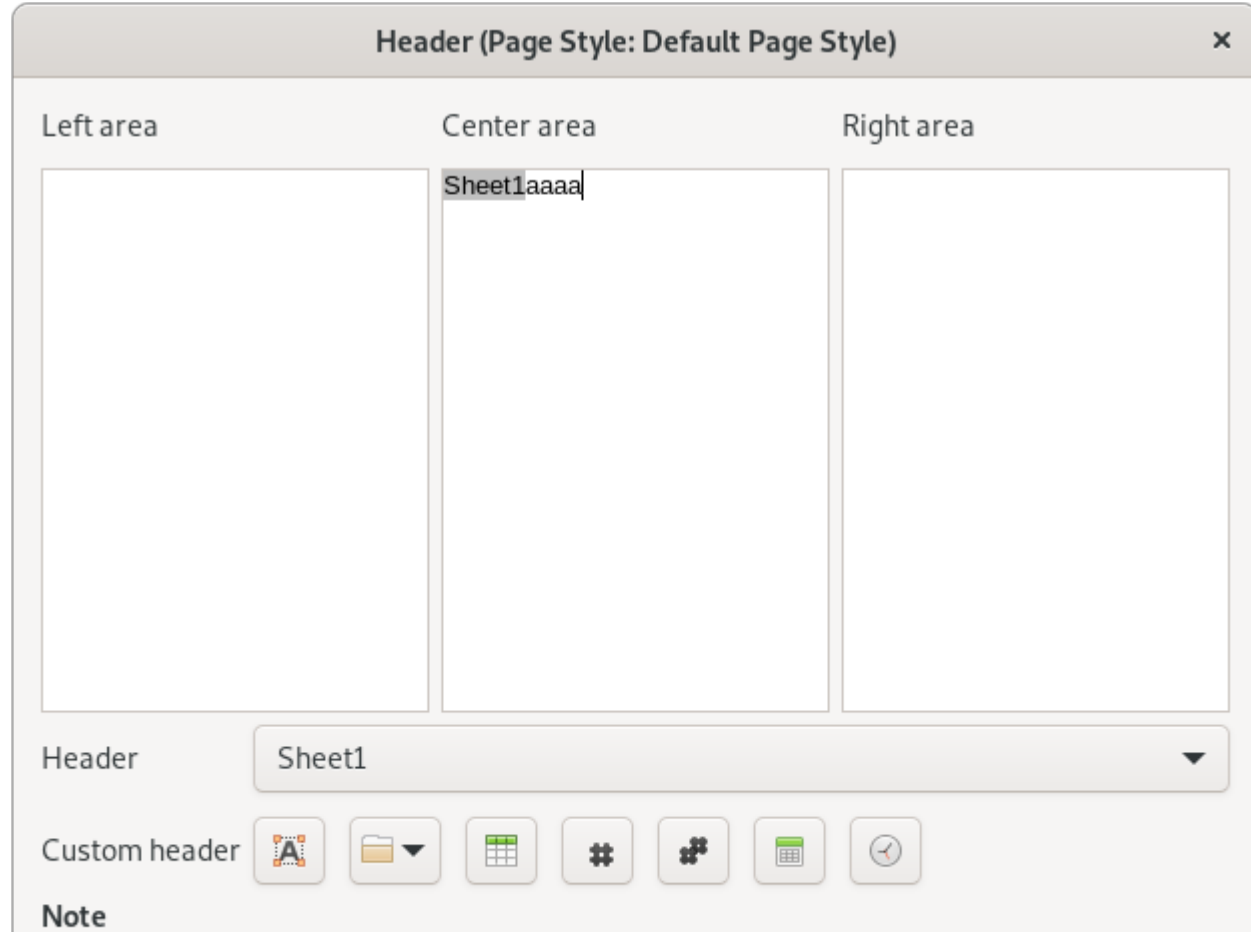There are many cases where bugs are encountered by combining vertical writing with other functions.

ある日の暮方の事である。一人の下人が、羅生門の下で雨やみを待っていた。広い門の下には、この男のほかに誰も居ない。ただ、所々丹塗の剝げた、大きな円柱に、蟋蟀が一匹とまっている。羅生門が、朱雀大路にある以上は、この男のほかにも、雨やみをする市女笠や揉烏帽子が、もう二三人はありそうなものである。それが、この男のほかには誰もいない。何故かと云うと、この二三年、京都には、地震とか辻風とか火事とか饑饉とか云う災がつづいて起った。そこで洛中のさびれ方は一通りではない。旧記によると、仏像や仏具を打砕いて、その丹がついたり、金銀の箔がついたりした木を、路ばたにつみ重ねて、薪の料に売っていたと云う事である。洛中がその始末であるから、羅生門の修理などは、元より誰も捨てて顧る者がなかった。するとその荒れ果てたのをよい事にして、狐狸が棲む。盗人が棲む。とうとうしまいには、引取り手のない死人を、この門へ持って来て、棄てて行くと云う習慣さえ出来た。そこで、日の目が見えなくなると、誰でも気味を悪るがって、この門の近所へは足ぶみをしない事になってしまったのである。その代りまた鴉がどこからか、たくさん集って来た。昼間見ると、その鴉が何羽となく輪を描いて、高い鴟尾のまわりを啼きながら、飛びまわっている。ことに門の上の空が、夕焼けであかくなる時には、それが胡麻をまいたようにはっきり見えた。鴉は、勿論、門の上にある死人の肉を、啄みに来るのである。――もっとも今日は、刻限が遅いせいか、一羽も見えない。ただ、所々、崩れかかった、そうしてその崩れ目に長い草のはえた石段の上に、鴉の糞が、点々と白くこびりついているのが見える。下人は七段ある石段の一番上の段に、洗いざらした紺の襖の尻を据えて、右の頰に出来た、大きな面皰を気にしながら、ぼんやり、雨のふるのを眺めていた。作者はさっき、「下人が雨やみを待っていた」と書いた。しかし、下人は雨がやんでも、格別どうしようと云う当てはない。ふだんなら、勿論、主人の家へ帰る可き筈である。所がその主人からは、四五日前に暇を出された。前にも書いたように、当時京都の町は一通りならず衰微していた。今この下人が、永年、使われていた主人から、暇を出されたのも、実はこの衰微の小さな余波にほかならない。だから「下人が雨やみを待っていた」と云うよりも「雨にふりこめられた下人が、行き所がなくて、途方にくれていた」と云う方が、適当である。その上、今日の空模様も少からず、この平安朝の下人の Sentimentalisme に影響した。申の刻下りからふり出した雨は、いまだに上るけしきがない。そこで、下人は、何をおいても差当り明日の暮しをどうにかしようとして――云わばどうにもならない事を、どうにかしようとして、とりとめもない考えをたどりながら、さっきから朱雀大路にふる雨の音を、聞くともなく聞いていた。

Can't input Japanese characters in Header and Footer Dialog: Bug#134566

We can't input any

Japanese characters

(through IME)

Caolán created a patch

Fixed in 7.0.3

Japanese Traditional Numeric texts are wrong such as 1,2,3, 10, 1000, 10000: Bug #130193

CJK Traditional Number text units are wrong in Calc

Naruhiko and DaeHyun created patch

himajin100000, Mark, and Ming commented

CJK contributors collaborated

Fiexd in 7.0

A	B [DBNum1][$-404]G/通用格式 Modern: mapped to [DBNum1]General in LibreOffice.	C [DBNum2][$-404]G/通用格式 Traditional: mapped to [DBNum2]General in LibreOffice.	D [DBNum3]0 I didn't see a fullwidth-arabic from UI of MSO tc 2013. I have to enter the format code manually.
0	○	零	0
1	一	壹	1
2	二	貳	2
3	三	參	3
4	四	肆	4
5	五	伍	5
6	六	陸	6
7	七	柒	7
8	八	捌	8
9	九	玖	9
10	一十	壹拾	1十
100	一百	壹佰	1百
1000	一千	壹仟	1千
10000	一萬	壹萬	1萬
100000000	一億	壹億	1億

Conclusion

- Many CJK functions are implemented in LibreOffice

- CJK function is often broken
  - Of course, mostly okay

- I will collect test cases and I will do test for CJK
  - I'm trying to increase the number of cases little by little

- Contact : shinji.enoki@libreoffice.org

Thank You!

ありがとうございました！