State of CJK issues of LibreOffice, 2019 edition

Shinji Enoki
shinji.enoki@gmail.com
ALMERIA | 13 Sept. 2019
Agenda

• What are CJK issues?
• Typical CJK functions
• CJK bug status
• Typical CJK issues
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
  - Organize member of Documentation Meetup
  - Staff of KANSAI OPEN SOURCE
Do you know the word “CJK issues”? 

• Please raise your hand if you don't know...
What are CJK issues / bugs

- CJK is an abbreviation for "Chinese-Japanese-Korean"
- LibreOffice has many language-specific features and issues, CJK issue is one of them.
- CJK issues happen when used in CJK's environment
  - A Function of CJK is broken
  - A bug occurs only with the CJK environment
Next Decade Manifesto

For the past ten years, the OpenOffice.org community has developed, supported and promoted the world's leading open-source office productivity suite. We have attracted the support of tens of thousands of individuals and corporate bodies during this period.

We now call on all our supporters to follow us into the next phase of our development, as we become an independent Foundation.

Our Values

WE COMMIT OURSELVES:

To eliminate the digital divide in society by giving everyone access to office productivity tools free of charge to enable them to participate as full citizens in the 21st century

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

WE REJECT:

The ownership of office productivity tools by monopoly suppliers which imposes a de-facto tax on global electronic free speech and penalises the economically disadvantaged

The creeping domination of computer desktops by a single language, forcing all people to learn a foreign language before they can express themselves electronically
"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
Why do CJK people need to work CJK bugs?

- It is important for CJK users to report / reproduce CJK bugs
- Because it is hard to notice / understand CJK bug unless it is people of CJK
Typical 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
• 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/

• And other language
What is “Requirements for Japanese Text Layout”

• According to Abstract
  - “This document describes requirements for general Japanese layout realized with technologies like CSS, SVG and XSL-FO.”
  - “The document is mainly based on a standard for Japanese layout, JIS X 4051, however, it also addresses areas which are not covered by JIS X 4051.”
• Currently, “Requirements for Japanese Text Layout” is collecting feedback for version upgrade.

• [https://github.com/w3c/jlreq](https://github.com/w3c/jlreq)
Vertical writing (1)

- Supported by Writer, Calc, Impress, Draw
  - Vertical text box
    - Note: Different from normal text box
    - Right click on the toolbar-> [Vertical Text]
• Writer: [Page Style] setting can be set to vertical writing
Phonetic guides (ruby)

- In Japanese and Chinese, mainly indicate how to read

https://www.w3.org/TR/jlreq/#ruby_and_emphasis_dots
- Writer: menu [Format] [Asian Phonetic guide]
• Line breaking rules “禁則処理” Kinsokushori
  - [Paragraph style]-[Japanese appearance] tab option
彼は背後にひそかな足音を聞いた。それはあまり良い意味を示すものではない。誰がこんな夜更けに、しかもこんな街灯のお粗末な港街の狭い小道で彼をつけて来るというのだ。人生の航路を捻じ曲げ、その獲物と共に立ち去ろうとしている、その丁度今。 彼のこの世事への恐れをして、待っている者がいるというのか。それとも背後の足音の主は、この街に無数にいる法監視役で、強硬な罰をすぐにも彼の手首にガシャンと下すというのか。
彼は足音が止まったことに気が着いた。あわてて辺りを見回す。ふと狭い抜け道に目が止まる。

OFF "Allow hanging punctuation"

彼は背後にひそかな足音を聞いた。それはあまり良い意味を示すものではない。誰がこんな夜更けに、しかもこんな街灯のお粗末な港街の狭い小道で彼をつけて来るというのだ。人生の航路を捻じ曲げ、その獲物と共に立ち去ろうとしている、その丁度今。 彼のこの世事への恐れをして、待っている者がいるというのか。それとも背後の足音の主は、この街に無数にいる法監視役で、強硬な罰をすぐにも彼の手首にガシャンと下すというのか。
彼は足音が止まったことに気が着いた。あわてて辺りを見回す。ふと狭い抜け道に目が止まる。

OFF "Allow hanging punctuation" and "Apply list of forbidden characters to the beginning and end of lines"
In Japan, both the Western calendar and Japanese calendar are used.

The Japanese calendar changes when the Emperor changes.

It was changed on May 1 this year.
Eike had committed a patch with dummy data in advance.

Abe-san committed in about 3 minutes after the announcement.

LibreOffice was able to respond very quickly.
CJK bug status
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
**Depends on CJK Meta issue #83066**

<table>
<thead>
<tr>
<th></th>
<th>2017/10</th>
<th>2018/9</th>
<th>2019/9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depends on:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(All)</td>
<td>135</td>
<td>168</td>
<td>186</td>
</tr>
<tr>
<td>Open</td>
<td>59</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>Close</td>
<td>76</td>
<td>95</td>
<td>113</td>
</tr>
</tbody>
</table>

Maybe some issues are not tracked by this issue.
Typical CJK issues
Alignment ignored when text grid:

**Bug #110994**

- Invalid alignment setting
- Mark Hung created a patch
- 6.0 : FIXED
Bug#111967

- There seems to be a crash problem
- Mark Hung created a patch
- 6.0 : FIXED
Noto Sans CJK fonts exported bug:
Bug #96091

- When exporting Noto Sans CJK font to PDF, paragraph decoration location is wrong
- Mark Hung created a patch
Noto Sans font's Some Old Hangul syllables have abnormal shown in exported PDFs: Bug#127010

- DaeHyun discovered this bug when I talked about #96091 at COSCUP 2019
- There are sometimes font problems like this
Mark Hung is fixing some CJK issues

- Thanks Mark!
  - Mark's keynote talk at LibreOffice Asia Conference
    - “LibreOffice CJK Bugs, Fixes, and Stories.”
  - Silide is blow:
    https://conf.libreoffice.jp/program.html
Font language settings change suddenly

- With Writer and Impress, Maybe other, all OS
- Problems from the OOo era
- I have not been able to identify the conditions to reproduce
• Many CJK functions are implemented in LibreOffice
• CJK function is often broken
• I will collect test cases and I will do test for CJK
Thank you!
Shinji Enoki (shinji.enoki@gmail.com)