LibreOffice
The Document Foundation

ROME CONFERENCE

State of CJK issues of LibreOffice

Shinji Enoki

shinji.enoki@gmail.com

ROME | 11 October 2017

Talk summary

- LibreOffice have many language-specific features and issues, CJK (Chinese / Japanese / Korean) issue is one of them.

- I will explain what is the CJK features and issues.

Who am I?

- Shinji Enoki（榎真治）
  - A member of LibreOffice Japanese Team(2011-)
    - Founding members
    - Join LibreOffice community in 2010
  - A member of TDF(2014-)
  - Activity: organize events, grows community, sometimes QA

LibreOffice Kaigi

What is CJK issues

- LibreOffice Bugzilla have the CJK meta issue

    - Bug 83066(CJK) - [META] CJK issues

    - CJK bugs are tracked META issues

ROME
CONFERENCE

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: Other All

Importance: medium normal
Assignee: Not Assigned
QA Contact:

URL:
Whiteboard:
Keywords:

Reported: 2014-08-25 14:50 UTC by Matthew Francis
Modified: 2017-10-20 17:03 UTC (History)
CC List: 14 users (show)

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

Crash report or crash signature:

Depends on: 35208 35301 36709 39935 43671 50349 Japanese-ruby-character 50901 55954 63011 65266 66791 69809 71329 72042 75790 76255 77803 80346 80812 81272 81484 82115 82116 82293 82689 82690 82692 82930 83844 85795 88752 89422 89979 90874 93695 95656 96197 101679 101895 103528 104378 104503 Text-Grid 107056 107195 *107383* 107435 107466 107467 *107484* 107485 107487 109030 109260 111717 112477 112593 112745 112879 *113048* 113184 *113189* 113191 CJK-Chinese-Traditional CJK-Chinese-Simplified CJK-Japanese CJK-Korean 40091 40335 44784 47719 49073 57139 58604 62750 64975 66126 67814 70097 76291 77076 77490 77514 77971 77972 78789 79766 80272 80473 80513 80604 80788 81144 82176 82818 83060 83145 83178 83252 83559 83752 83755 84138 85954 87626 88169 88527 88855 89409 91288 92655 92782 94008 94066 95379 95836 96091 96457 96829 97405 97702 98165 98269 98527 98879 101729 102041 102074 102253 103710 103718 103785 103850 104625 104726 104801 104828 104954 106097 106295 107462 112066 112807

Blocks:

Show dependency tree / graph

Attachments
Add an attachment (proposed patch, testcase, etc.)

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 this "Our Values" is very good

- Our problem want to resolve ourselves although it's not easy to do

ROME
CONFERENCE

Last year's Conference: Talk

- Kei Tasaki talked "Typical" bugs of LibreOffice in Japanese environment

  - How to create a document in Japan

  - "Typical" bugs in Japanese environment

Issues status of Last year talk

- Ruby using case bug

  - Bug 98269 - FORMATTING: Different spacing with Japanese ruby text between WIN and LINUX

    - <u>RESOLVED DUPLICATE</u> → <u>VERIFIED FIXED</u>

  - Bug 77514 - FORMATTING: CJK ruby text (furigana) in vertical mode pushes characters to the left of the base line

    - <u>RESOLVED NOTABUG</u>

フリーソフトウェア (free software) とは、ソフトウェアのうち、フリーソフトウェア 財団（ざいだん）が提唱する自由（じゆう）なソフトウェアを指す。大半のフリーソフトウェアは無償（フリー）として配布されているが、定義に従えば、現状事実上無償での入手がほぼ可能とみてよいが、必ずしも全ての一次配布が無償ということは定義にはない。

財団（ざいだん）　○
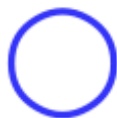
財団（ざいだん）　×
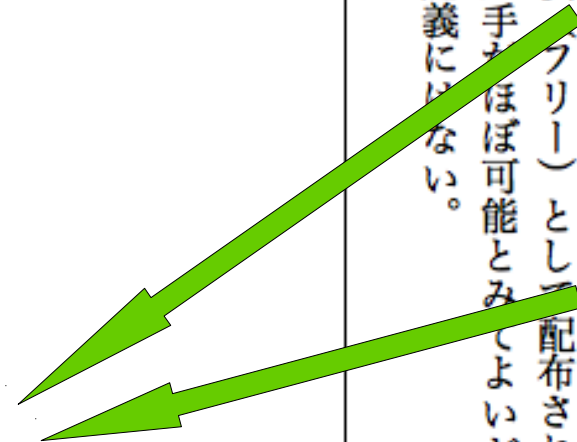
Bug 77514

- We can work around by setting that Mark Han advice

- I think that it is better to have such a setting by default in the Japanese environment.
  How about in other languages?

- Should we make "Request for Enhancement" as issue ? That is , Default setting switches according to languages

- Reference about ruby :
  https://www.w3.org/TR/jlreq/#ruby_and_emphasis_dots

External Character（異体字）
& Kanji variants

Kanji variant is a Kanji character with the same meaning but refers to what the shape is different

崎 﨑 埼
碕 嵜

All character is same means.

IVS(Ideographic Variation Sequence)

In Japan, using a mechanism called IVS to make it easier to deal with a Kanji variant.

IVS is conbine the characters, as one of the character.

Unicode character

葛

Variation Sequence

葛

UCSコードは、細かな字形の差異は区別せず、複数の字形を一つのコードに統合している。

UCSコード

UCSコードと字形を指定するVS（字形選択子）を並べて記述することで、字形を指定することができる。※VSの値は実際の規格と異なります。

VS

甚 甚 …　845A

葛 葛 葛 …　845B　葛 E0100
　　　　　　　　葛 E0101
　　　　　　　　葛 E0102

辺 辺 …　8FBA

辻 辻 …　8FBB　辻 E0100
　　　　　　　　辻 E0101

http://mojikiban.ipa.go.jp/1292.html

ROME CONFERENCE

14

IVS using case bug

- This issue has not reported yet
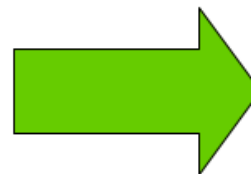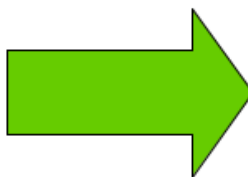
# IVS problem is not fixed
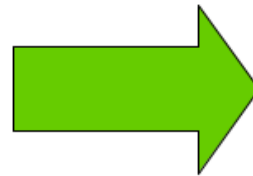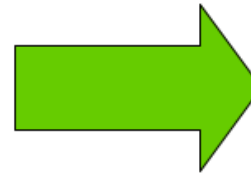
Push Backspace key

CJK features

- In common with CJK, 1.use Kanji (Chinese characters), 2. vertical writing, 3.phonetic guides (ruby)

- Chinese characters are also different in Simplified Chinese (Mainland China), Traditional Chinese (Taiwan, Hong Kong) and Kanji (Japanese)

  - Entering kanji in this template turned to Simplified Chinese

- Even with common functions, usage methods are not always the same

  - Vertical writing

ROME
CONFERENCE

Features specific to Japanese

- Line breaking rules 禁則処理 Kinsokushori

  - [Paragraph style]-[Japanese appearance] tab option

ROME CONFEREN

18
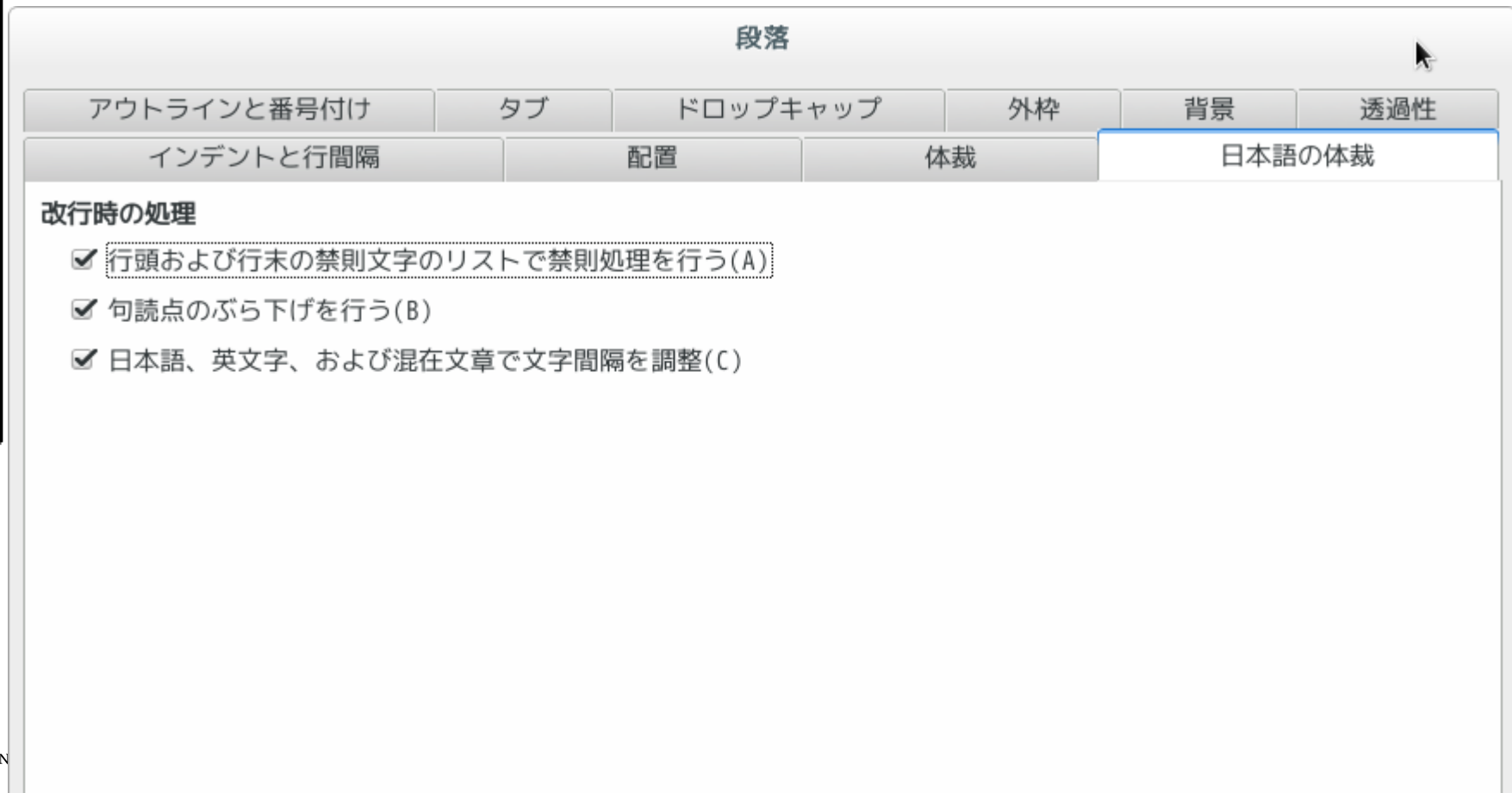
Check [META] CJK issues

Depends on:(All)	135
Open	59
Close	76

- Mark Hung,is fixing CJK ISSUE Thank you!

- Some cases is not added to CJK issue due to Japanese related bug. I will add it in the future

Checking is difficult

- It is difficult to understand the meaning

- Test case, sample file is not enough

- I do not understand when I read the comment to the end

- I wanted to re-make or arrange the summary and test cases

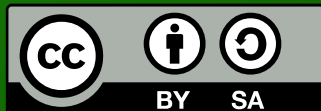# Multilingual issues that matter even at CJK

- Multilingual problem that becomes a problem with CJK

- In some time zone Date is shifted

  - Bug 59850 - EDITING: typed date decreased one day

  - Bug 63230 - Calc: Date decreased by 1 depending on Timezone

Conclusion

- I share the Unique function and issue about CSK with a lot of members

- I'd like to work out to solve issues with our Japanese members

- I want to progress organization of Bugzilla such as CSK issue

Thank you !

shinji.enoki@gmail.com

ROME
CONFERENCE