



LibreOffice
The Document Foundation



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



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



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



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/>
- 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. "



Jlreq: upgrade work in progress

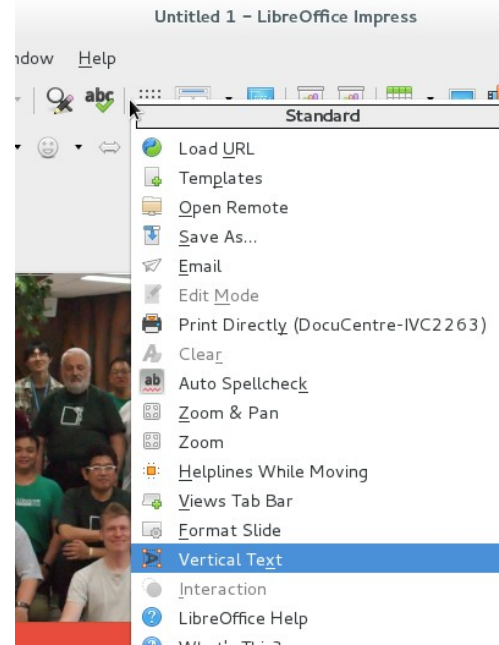
- Currently, "Requirements for Japanese Text Layout" is collecting feedback for version upgrade.
- <https://github.com/w3c/jlreq>



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]

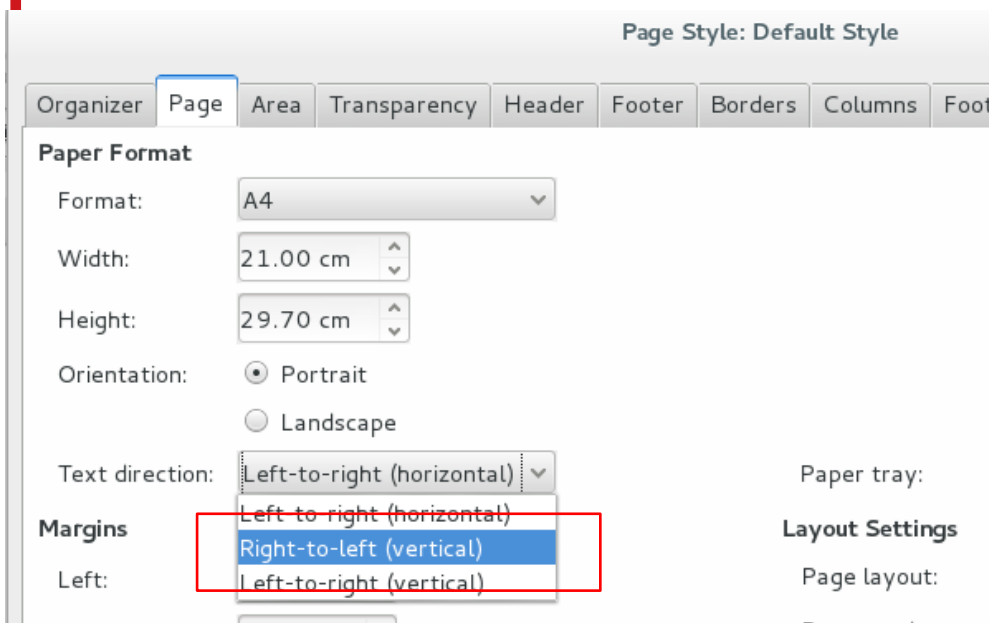
縦書きの
Vertical text
box





Vertical writing (2)

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



俊寛様の話ですか？ 俊寛様の話くらい
たし、——自身の事さえ、でもない嘘が
いをなすってしまおうし、わたしはそのを
あの島の女と、夫婦のらいをなすったと
した。前の琵琶法師の語った事が、もた
やはりいい加減の出たらめなのです。
一体琵琶法師などと云うものは、どわ
わたしはあの小屋に、俊寛様が子供を
つい涙さえ落しました。たとい嘘とは三
があるだけ、わたしでも今の内ありのま
こうあなたはおっしゃるのですか？ な
の事を御話しましょう。しかしわたしは
ない真実と云う事だけです。ではどうか





Phonetic guides (ruby)

- In Japanese and Chinese, mainly indicate how to read

base character → 君 く ← ruby
 → 君 ん ← ruby

base character → 子 し ← ruby

base character → 和 わ ← ruby

base character → 同 どう ← ruby
 → 同 とう ← ruby

ず

くん し わ どう
君子は和して同ぜず



- Writer: menu [Format] [Asian Phonetic guide]

の か で 見 め 右 に 仲 っ 彼
間 け は 背 後
の て 来 後
中 来 後
に る に
同 と ひ
じ い そ
考 う か
え の な
を だ 足
抱 き 音
、 人 を
生 聞
彼 の いた
を 航 路
見 路 を
守 り 捻
、 じ は
待 曲 げ
っ げ ま
て っ り
い そ 良

Asian Phonetic Guide

Base text	Ruby text
背後	はいご

Alignment: 0 1 0 ▾ Position: Top ▾ Character style for ruby text: Rubies ▾ Styles

Preview:

はいご

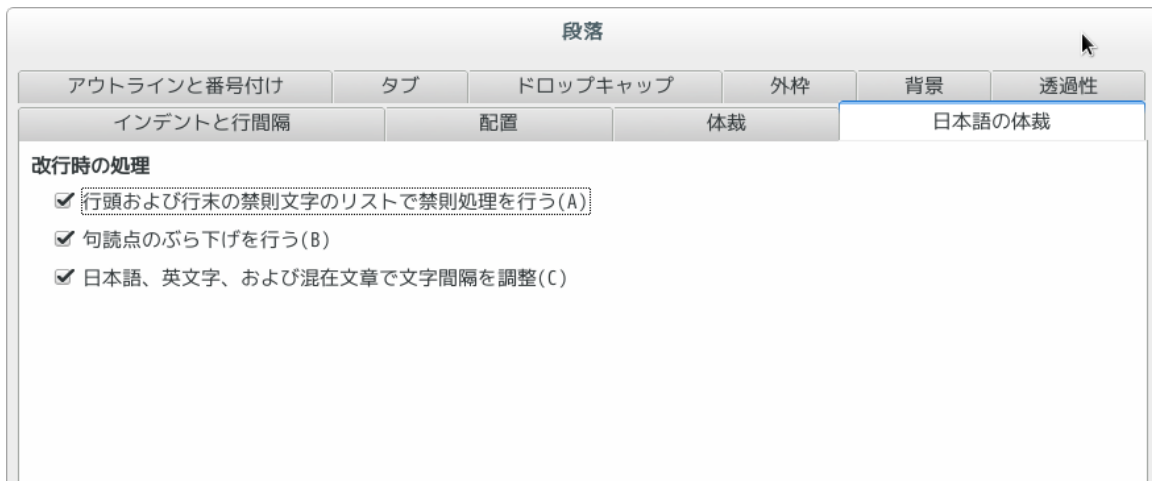
背後

Help Close Apply



Line composition: example

- Line breaking rules “禁則処理” Kinsokushori
 - [Paragraph style]-[Japanese appearance] tab option





Line composition: example (2)

はいこ
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なのだ。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。彼のこの仕事への恐れを和らげるために、数多い仲間の中に同じ考えを抱き、彼を見守り、待っている者がいるといなのだ。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、強靱な罰をすぐにも彼の手首にガシャンと下すといなのだ。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"

はいこ
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なのだ。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。彼のこの仕事への恐れを和らげるために、数多い仲間の中に同じ考えを抱き、彼を見守り、待っている者がいるといなのだ。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、強靱な罰をすぐにも彼の手首にガシャンと下すといなのだ。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"

はいこ
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なのだ。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。彼のこの仕事への恐れを和らげるために、数多い仲間の中に同じ考えを抱き、彼を見守り、待っている者がいるといなのだ。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、強靱な罰をすぐにも彼の手首にガシャンと下すといなのだ。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。



Japanese calendar : Calc

	A	B
1	Westen	Japanese
2	2019/04/30	平成31年4月30日
3	2019/05/01	令和1年5月1日

- 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



Japanese calendar : Calc (2)

- 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



index : libreoffice/core

main, development code repository

[summary](#) [refs](#) [log](#) [tree](#) **[commit](#)** [diff](#)

author Takeshi Abe <tabe@fixedpoint.jp> 2019-04-01 11:44:32 +0900
committer Eike Rathke <erack@redhat.com> 2019-04-02 22:54:49 +0200
commit [597c5d75b8e72d429e096535334eaac7973455ef](#) (patch)
tree [dc3ee3d8645dfa28619e5c7cd1fd66ef50eec548](#)
parent [66c08a0d308fd6119460546a906434bbc9d23c2e](#) (diff)

Introduce next Japanese gengou era 'Reiwa'

starting from 2019-05-01, which has been announced officially.

This fills the provisional slot acknowledged at [cacbb0faef77ae8462de9ff5c7307a6a2e28b2bb](#).

Change-Id: Ifb12e6afaad4c66d455f664b46ec946e80324e87
Reviewed-on: <https://gerrit.libreoffice.org/70028>
Tested-by: Jenkins
Reviewed-by: Eike Rathke <erack@redhat.com>



CJK bug status

Bug 83066 (CJK) - [META] CJK (Chinese, Japanese, Korean, and Vietnamese) language issues**Status:** NEW**Alias:** CJK**Product:** LibreOffice**Component:** LibreOffice ([show other bugs](#))**Version:** unspecified
(earliest affected)**Hardware:** Other 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](#) [85795](#) [88752](#) [89422](#) [89979](#) [90874](#) [93695](#) [94631](#) [101679](#) [101895](#) [104378](#) [104503](#) [104927](#) [105501](#) [107195](#) [107466](#) [107467](#) [107484](#) [107485](#) [108042](#) [109030](#) [113048](#) [113184](#) [113189](#) [113191](#) [114272](#) [114520](#) [114746](#) [114760](#) [114761](#) [114763](#) [114764](#) [115321](#) [115970](#) [117554](#) [118260](#) [118476](#) [118683](#) [118849](#) [119219](#) [120171](#) [121527](#) [125404](#) [125446](#) [126157](#) [126169](#) [126406](#) [127010](#) [127491](#) [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](#) [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](#) [121855](#) [122459](#) [123956](#) [123957](#) [124657](#) [125620](#)

Blocks: [Languages](#)Show dependency [tree](#) / [graph](#)**Reported:** 2014-08-25
14:50 UTC by
Matthew
Francis**Modified:** 2019-09-11
07:29 UTC
([History](#))**CC List:** 16 users
([show](#))**See Also:** [Font-Rendering](#)
[Vertical-Text](#)
[RTL-CTL](#)
[Ruby](#)
[Text-Grid](#)
[CJK-Chinese-Traditional](#)
[CJK-Chinese-Simplified](#)
[CJK-Japanese](#)
[CJK-Korean](#)**[Crash report or crash signature:](#)**



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

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

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

5.4

]Text-to-text Alignment: baseline				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: top				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: center				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: bottom				

6.3

]Text-to-text Alignment: baseline				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: top				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: center				
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
今	日	の	会	議
Text-to-text Alignment: bottom				



Vertical text in table formatted incorrectly: Bug#111967

- There seems to be a crash problem
- Mark Hung created a patch
- 6.0 : FIXED

About LibreOfficeDev

LibreOffice

Version: 6.0.0.0.alpha0+
Build ID: cd65bae41c52c4d2f40589776a645e224ee222cd
CPU threads: 2; OS: Linux 4.4; UI render: default; VCL: gtk2;
Locale: zh-TW (zh TW.UTF-8); Calc: group

[See Log: cd65bae41c52c4d2f40589776a645e224ee222cd](#)

LibreOfficeDev is a modern, easy-to-use, open source productivity suite for word processing, spreadsheets, presentations and more.

This release was supplied by markhung.
Copyright © 2000–2017 LibreOffice contributors.
LibreOfficeDev is derived from LibreOffice which was based on OpenOffice.org.

Credits Website Close

LibreOffice について

LibreOffice
The Document Foundation

バージョン: 6.2.6.2
Build ID: 684e730861356e74889dfe6dbddd3562aae2e6ad
CPU threads: 4; OS: Linux 3.16; UI render: default; VCL: gtk2;
ロケール: ja-JP (ja_JP.utf8); UIの言語: ja-JP
Calc: threaded



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

PDF created with LO 5.1.1.2

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。

Evince 3.14.2

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。

Fixed

PDF.js on Fx 45

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。

much better, but still a little bit too high

Chromium 48

Droid Sans Fallback

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色

Noto Sans CJK TC Regular

觀自在菩薩，行深般若波羅蜜多時。照見五蘊皆空，
渡一切苦厄。ABCED TEST 舍利子，色不異空，空不
異色，色即是空，空即是色。



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

ODT

옛 한글 테스트

골 가람 강(江)

들 달

제주어(濟州語, Jeju-eo)

흔저 옴서예 - 안녕하세요

뭉국 - 뭉국

月彔印인千천江강

PDF

옛 한글 테스트

골 가람 강(江)

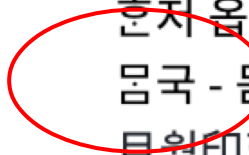
들 달

제주어(濟州語, Jeju-eo)

흔저 옴서예 - 안녕하세요

뭉궂 - 뭉궂

月彔印인千천江강





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:

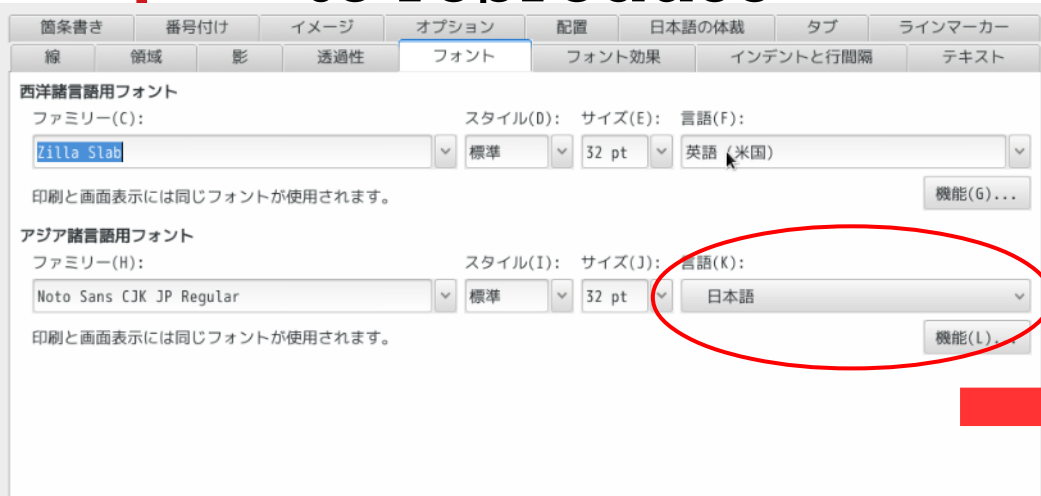
<https://conf.libreoffice.jp/program.html>





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





Conclusion

- 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)



All text and image content in this document is licensed under the Creative Commons Attribution-Share Alike 4.0 License (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these thereof is subject to trademark policy.