

日本語組版処理の要件の現状

Atsushi Shimono (W3C/Keio)

Outline

W3Cの国際化ワーキンググループでは、すべての言語を利用可能にすることを目標に、言語固有機能・要求と仕様・実装との齟齬解析を行い仕様開発やブラウザにフィードバックすることや、言語毎の組版処理の要件をまとめた文書の開発などを行っています。このトークでは

- 国際化ワーキンググループ概要
- 日本語組版処理タスクフォースのこれまでの活動
- リフロー前提、かつUnicodeを前提としたデジタル時代のJLReqに向けて

についてご紹介します。

W3C国際化ワーキンググループの活動

CSSなどを含め、Web仕様は日々更新されつつあり、それらに対して様々な言語に対して対応しているかのチェックを随時入れることが重要となっています。

また、よりきちんと様々な言語をWeb上で表現できるようにするために足りない機能を仕様化し実装されるようにすることも必要となります。

Phase 1 – 言語ごとの要求仕様のとりまとめ

Phase 2 – 実現しているWeb仕様との齟齬解析

Phase 3 – 言語話者と仕様開発者のネットワーク

これらに加え、仕様開発における、国際化に関する参考文書や自己レビューリストなども開発し仕様開発者向けなどに提供しています。

xLReq - 組版要件

5言語分が公開されている

一つの言語についてそれなりの人数の言語専門家が存在するか、国家標準的なまとまった文書が存在するとまとめられやすい、という規模の文書ではある。

Requirements for Japanese Text Layout

W3C Editor's Draft 12 March 2019



This version:

<https://w3c.github.io/jlreq/>

Latest published version:

<https://www.w3.org/TR/jlreq/>

Latest editor's draft:

<https://w3c.github.io/jlreq/>

Bug tracker:

[File a bug](#) ([open bugs](#))

Editors:

阿南 康宏 (Yasuhiro Anan), (第1版 1st edition) (Microsoft)
千葉 弘幸 (Hiroyuki Chiba), (第2版 2nd edition) (Invited Expert)
枝本 順三郎 (Junsaburo Edamoto) (Invited Expert)
Richard Ishida (W3C)
石野 恵一郎 (Keiichiro Ishino) (Antenna House)
加藤 誠一 (Seiichi Kato), (第2版 2nd edition) (Microsoft)
小林 龍生 (Tatsuo Kobayashi) (JustSystems)
小林 敏 (Toshi Kobayashi) (Invited Expert)
小野澤 賢三 (Kenzou Onozawa) (Invited Expert)
Felix Sasaki (DFKI GmbH)
塩澤 元 (Hajime Shiozawa), (第2版 2nd edition) (Invited Expert)

Github:

[repository](#)

国際化に関する参考文書

仕様を策定するうえで参考になるようなドキュメントを作成、仕様策定者に提示するとともに、W3Cの水平レビューの枠組みの中で全ての仕様を国際化の観点からレビューしています。

- [Strings on the Web: Language and Direction Metadata](#)
- [International text layout and typography index](#)
- [Character Model for the World Wide Web: String Matching](#)
- 仕様のチェックリスト : [shortchecklist](#), [i18n techniques](#)

“JLreq”に関連する歴史

1995 JIS X 4051:1995

2004 JIS X 4051:2004

2006 W3C Print Symposium

2007 W3C Japanese Layout Task Force (JL-TF)開始

2008 “Requirements of Japanese Text Layout” (JLreq) draft公開

2009 JLreq WG Note ver.1公開

2012 JLreq WG Note ver.2公開

2013 “Requirements for Hangul Text Layout and Typography” (KLReq) Working Draft (FPWD)

2014 “Requirements for Latin Text Layout and Pagination”公開

2015 “Requirements for Chinese Text Layout” (CLReq) Working Draft (FPWD)

JLreqの概要

JIS X 4051をベースとして、冊子などの紙媒体で行われてきた(活版や写植の時代からの歴史的な構造などを含む)版面構成や組版規則、そして考慮事項をまとめた文書として英語・日本語両方で提供しています。

日本語とその文字などについての知識がない仕様開発者にとっても把握可能とするために、さまざまな基礎的事項も追加されています。

目次

Abstract

1. 序論

- 1.1 この文書の目的
- 1.2 この文書の作成方法
- 1.3 この文書の執筆方針
- 1.4 この文書の構成
- 1.5 用語の参照その他

2. 日本語組版の基本

- 2.1 日本語組版に使用する文字と配置の原則
- 2.2 日本語文書の基本となる組体裁
- 2.3 組方向(縦組と横組)
- 2.4 基本版面の設計
- 2.5 基本版面の設計要素の各ページに対する適用
- 2.6 柱とノンブル

3. 行の組版処理

- 3.1 約物などの組版処理
- 3.2 和欧文混植処理(縦中横処理を含む)
- 3.3 ルビと圈点処理
- 3.4 割注処理
- 3.5 段落整形, そろえ及び段落末尾処理
- 3.6 タブ処理
- 3.7 その他の行組版処理
- 3.8 行の調整処理
- 3.9 文字クラスについて

4. 見出し・注・図版・表・段落の配置処理

- 4.1 見出し処理(改ページ処理も含む)
- 4.2 注の処理
- 4.3 図版の配置処理
- 4.4 表の処理
- 4.5 行・段落などの行送り方向の配置処理

A. 文字クラス一覧

A. 文字クラス一覧

B. 文字間の空き量

C. 文字間での分割の可否

D. 行の調整処理で詰める処理が可能な箇所

E. 行の調整処理で空ける処理が可能な箇所

F. 熟語ルビの配置方法

G. 用語集

H. 参考文献

I. 謝辞

J. 変更記録

2018年、JL-TF再始動

CSSにも縦書きに関連するさまざまなプロパティが入ったりと、ここ10年で日本語組版で必要とされる機能がいくつか実装されてきました。また、文字コードについてもJISからUnicodeが基盤となるなど、合わせて利用される周辺の標準技術の環境にも大きな変化が起こりました。

この環境の中、いまだ長らく残っている課題点に対応することと、標準技術に追随した更新を加えることなどを目標にして活動しています。

- JLreqの文書の次期バージョンの開発
- 重点の一つの<ruby>に関する文書の策定
- css-textなどのCSS仕様に対してのインプット

<ruby>にまつわる諸問題

<https://html.spec.whatwg.org/multipage/obsolete.html#non-conforming-features>

<ruby><rb>日本<rt>にほん<rb>語<rt>ご<rb>組<rt>くみ<rb>版<rt>はん</ruby>

<ruby><rb>日本<rb>語<rt>にほん<rt>ご<rb>組<rb>版<rt>くみ<rt>はん</ruby>

16.2 Non-conforming features

Elements in the following list are entirely obsolete, and must not be used by authors:

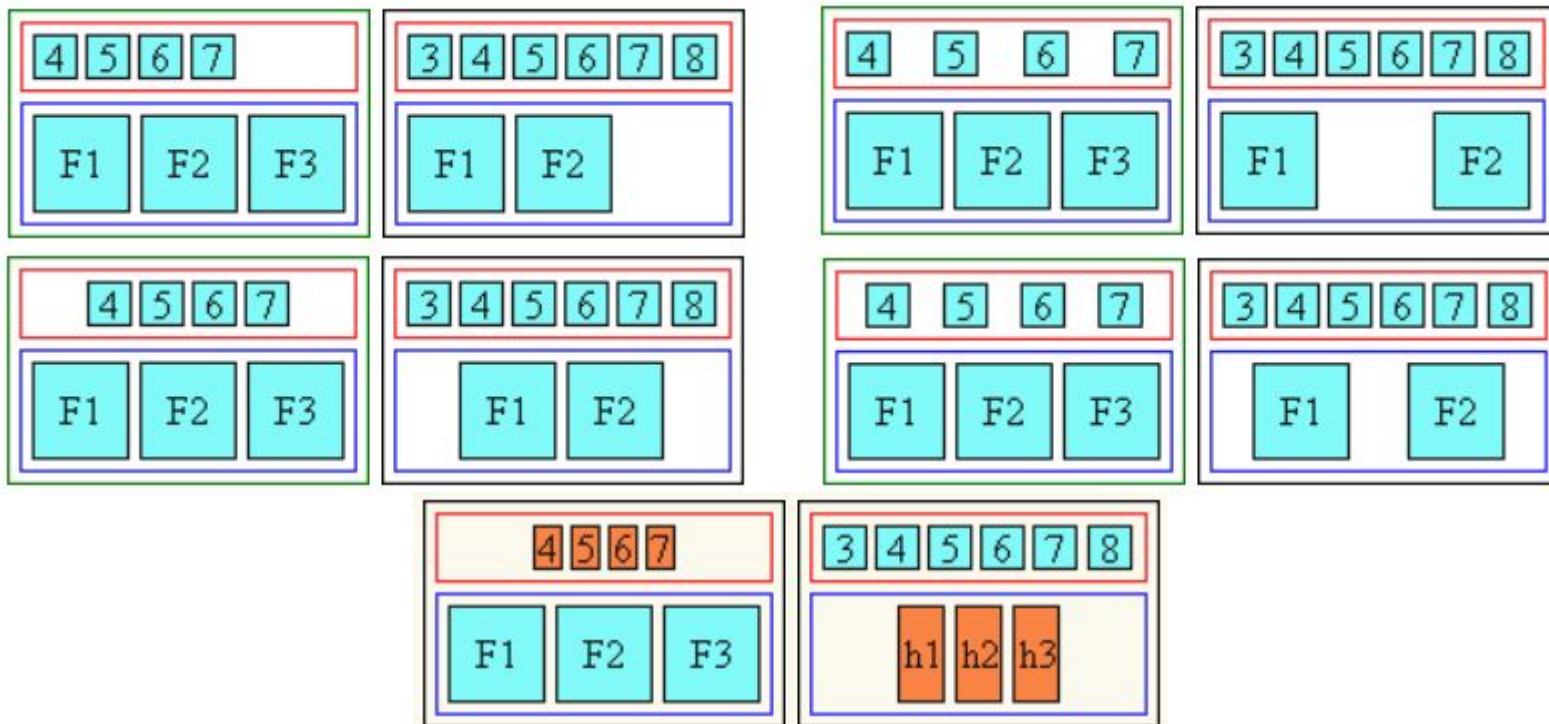
rb

rtc

Providing the ruby base directly inside the **ruby** element or using nested **ruby** elements is sufficient.

<ruby>にまつわる諸問題

<https://www.w3.org/TR/css-ruby-1/>



ルビの簡便な配置ルール (simple-ruby)

ルビ文字の配置処理について、JLreqでは活版の時代から綿々と開発され続けている例外が大量にある配置ルールを可能な限り記載するという方針で、さまざまな例外処理を含めて詳細に記述されていますが、実際に仕様化するのは困難を伴います。

また、漢字に平仮名を振るというのではない、CJK文字以外を利用するルビなど、国際化による多様な事例にも対応が必要になったりします。

Rules for Simple Placement of Japanese Ruby

W3C Editor's Draft 17 March 2022



► **More details about this document**

This document is also available in this non-normative format: [Original Japanese \(PDF\)](#)

Copyright © 2022 W3C® (MIT, ERCIM, Keio, Beihang). W3C [liability](#), [trademark](#) and [permissive document license](#) rules apply.

Abstract

A simple set of rules for placement of Ruby text in Japanese typography.

ルビのアクセシビリティ考慮

ルビは親文字をどう読むかという情報を提示するためのモノ、というのが最もオーソドックスな利用方法ではありますが、その中においても全ての漢字にルビを振ったり、初出のみに振ったり、などさまざまな形態があります。また、辞書的な直接の読みではなく特定のコンテキストでのみ利用されるルビが振られることもあります。

これらについて、読み上げソフトでどう対応するかの考慮が必要です。

Text to Speech of Electronic Documents Containing Ruby: User Requirements

W3C Editor's Draft 03 March 2022

► [More details about this document](#)

[Copyright](#) © 2022 W3C® ([MIT](#), [ERCIM](#), [Keio](#), [Beihang](#)). [W3C liability, trademark and permissive document license rules apply.](#)

Abstract

This document describes user requirements for text to speech of electronic documents containing ruby.



JLreq-D - リフロー前提の電子媒体向け組版要件

from: <https://github.com/w3c/jlreq-d/blob/gh-pages/README.md>

ちょうど手書きによる書写と印刷のように、印刷とWebに代表されるデジタルテキストでは、技術的な成り立ちが根底から異なっている。一つの根本的な違いは誰が組版を決定するかにある。印刷においては送り手が、しばしばプロによる校正ののち組版を決定し、単一のイメージに固定したものを受け手に配布する。デジタルテキストにおいては組版は送り手の元で完成せず、受け手の元で、デバイスの特性や受け手の好みに従って無数の異なるイメージが生成される。このような違いが印刷と本質的に異なる制約条件と可能性を生み出す。

このため、印刷のためではない、デジタルネイティブな組版について記述した文書を開発することは有用であると考え。デジタルテキストにフォーカスした要件文章を JLreq の先にあるものとして開発することの合意が JLreq TF で行われた(2021/10)。とりあえずの略称を JLreq-d とする。

JLreq-D - リフロー前提の電子媒体向け組版要件

要件検討の基本方針として:

- Unicodeのように国際化ソフトウェア環境の中の一つの言語としての日本語
- 確立されている技術を基盤に: Unicode、ウェブ技術、高機能フォント、など
- シンプル化、優先順位付け、既定方式、などの提案により実装コストを低減
- フォントなど既存JLreqが扱っていない必要技術について日本語の特性を記述
- アクセシビリティを含むユーザによる最適な表示形態の選択を可能に

なお、日本語組版に精通していない技術者向けに基本部分から記述するなどの、これまでの記述方針は踏襲することとしています。

JLreq-D - リフロー前提の電子媒体向け経

現行のJLreqではJISの文字集合を基本として、配置処理の中で利用するカテゴリ分類を「文字クラス」として記述しています。また、この中にはルビや連数字などのブロックを仮想的な文字としてとらえたものも含まれます。

Unicodeでは各コードポイントについてさまざまな属性が付与されており、漢字・仮名などのカテゴリ分類から、始め括弧類などの意味的な分類などもあり、次期JLreq-dではUnicodeのさまざまな属性(のみ)を利用した文字クラスを定義することで、各種の仕様で利用しやすい形を目標とします。

A.	文字クラス一覧
A.1	始め括弧類 (cl-01)
A.2	終わり括弧類 (cl-02)
A.3	ハイフン類 (cl-03)
A.4	区切り約物 (cl-04)
A.5	中点類 (cl-05)
A.6	句点類 (cl-06)
A.7	読点類 (cl-07)
A.8	分離禁止文字 (cl-08)
A.9	繰返し記号 (cl-09)
A.10	長音記号 (cl-10)
A.11	小書きの仮名 (cl-11)
A.12	前置省略記号 (cl-12)
A.13	後置省略記号 (cl-13)
A.14	和字間隔 (cl-14)
A.15	平仮名 (cl-15)
A.16	片仮名 (cl-16)
A.17	等号類 (cl-17)
A.18	演算記号 (cl-18)
A.19	漢字等 (漢字以外の例) (cl-19)
A.20	合印中の文字 (cl-20)
A.21	親文字群中の文字 (添え字付き) (cl-21)
A.22	親文字群中の文字 (熟語ルビ以外のルビ付き) (cl-22)
A.23	親文字群中の文字 (熟語ルビ付き) (cl-23)
A.24	連数字中の文字 (cl-24)
A.25	単位記号中の文字 (cl-25)
A.26	欧文間隔 (cl-26)
A.27	欧文用文字 (cl-27)
A.28	割注始め括弧類 (cl-28)
A.29	割注終わり括弧類 (cl-29)
A.30	縦中横中の文字 (cl-30)

JLreq-D - リフロー前提の電子媒体向け組版要件

このほかにも課題として:

- 和欧文字間の空き量をどうするか
- 長いグループルビの場合のその間での改行
- Unicodeのプロパティとフォント(OTF/AJ1xなど)との関連
- OpenTypeフォントのテーブルと空き・詰めの方式との関連

などについても議論が繰り広げられています。必要に応じて、W3Cの国際化WGに閉じず、Unicode ConsortiumやCITPCなどさまざまなSDOsと連携しての議論を行っています。

Questions?