

音声読み上げによるアクセシビリティに対応した  
電子書籍制作ガイドライン

平成 27 年 6 月

総務省



## 用語説明

本ガイドラインで用いる用語の意味及び定義は、以下のとおりである。

名称	意味（解説及び定義）
TTS	Text to Speech の略。音声合成による読み上げのこと。音声読み上げ機能を示す場合は TTS 機能、音声読み上げソフトウェアの場合は、TTS ソフトと示す。
W3C	World Wide Web で使用される各種技術の標準化に賛同する企業、研究所、大学、政府機関、ユーザ団体などが集まり、標準化を推進するために 1994 年に設立された団体。日本では慶応義塾大学 SFC 研究所がホスト機関として活動している。
IDPF	International Digital Publishing Forum のことを示し、電子書籍ファイルフォーマット EPUB の策定と普及を推進している団体である。 <a href="http://www.idpf.org/">http://www.idpf.org/</a>
EPUB	IDPF が主体となって仕様の策定及び普及を行っている、電子書籍のファイルフォーマット。現在のバージョンは 3.0.1（表記上は EPUB3）。EPUB フォーマットは、個人用途、商用問わず、誰でも、特定の団体や個人に許諾を得る必要なく、自由に、無料で利用できる。なお本報告書で EPUB と表記した場合は、EPUB3 を表す。
DAISY	Digital Accessible Information SYstem（アクセシブルな情報システム）の略で、デイジーと読む。音声にテキスト及び画像を同期させることで、利用者は音声を聞きながらハイライトされたテキストを読むことができる（画像は位置の確認）。視覚障害や学習障害(LD)等により印刷物を読むことが困難な人々のために開発されたシステムである。 <a href="http://www.daisy.org/">http://www.daisy.org/</a>
SMIL	Synchronized Multimedia Integration Language の略で、W3C で規定されている言語仕様。スマイルと読む。静止画、動画、音声、テキストなどの、位置レイアウト、時間軸上でのレイアウトを、XML フォーマットで記述することで統合し、再生させることができる言語として規定されている。 <a href="http://www.w3.org/TR/SMIL/">http://www.w3.org/TR/SMIL/</a>

名称	意味（解説及び定義）
Media Overlays	<p>EPUB の仕様の一つ。主に、SMIL を利用してマルチメディアデータとテキストとの同期を取る仕掛けを示す。メディアオーバーレイと読む。</p> <p><a href="http://www.idpf.org/epub3/latest/mediaoverlays">http://www.idpf.org/epub3/latest/mediaoverlays</a></p>
SSML	<p>Speech Synthesis Markup Language の略で、W3C で規定されている言語仕様。日本語では、“音声合成マークアップ言語”という。合成音声を生成するのに必要な発音、声量、音高等々をテキストに対してマーク付けを行う目的の XML ベースの言語を示す。</p> <p><a href="http://www.w3.org/TR/speech-synthesis/">http://www.w3.org/TR/speech-synthesis/</a></p>
PLS	<p>Pronunciation Lexicon Specification の略で、W3C で規定されている言語仕様であり、発音辞書を規定するために用いられる。日本語では“発音辞書仕様”という。SSML に組み込むことができる発音辞書は PLS を利用して記述される。</p> <p><a href="http://www.w3.org/TR/pronunciation-lexicon/">http://www.w3.org/TR/pronunciation-lexicon/</a></p>

# 目次

用語説明.....	1
1. 読み上げ対応のための電子書籍記述仕様 .....	1
1.1. 音声読み上げのために必要なメタデータ .....	1
1.1.1. lang 属性： .....	1
1.1.2. gender 属性： .....	2
1.2. 音声読み上げ指定技術の用途と適用方法 .....	2
1.2.1. 全文 SSML と部分 SSML の用途.....	2
1.2.2. SSML の適用方法.....	3
2. アクセシビリティ対応電子書籍リーダー設計指針 .....	5
2.1. SSML 及び PLS の表記方法.....	5
2.1.1. 音声表記について .....	5
2.1.2. SSML の表記方法.....	5
2.1.3. PLS の表記方法 .....	6
2.1.4. 間の制御方法.....	7
2.1.5. SSML 及び PLS の本文への埋め込み方法.....	7
2.2. リーダー側に求められる機能 .....	7
2.2.1. リーダーの基本機能.....	7
2.2.2. 音声データと文字との同期機能 .....	8
2.2.3. 複数の本を同時に開く機能.....	8
2.2.4. 発音辞書ファイル（PLS）の登録及び解除機能.....	8
2.2.5. SSML データの受け渡し機能.....	9
2.2.6. PLS に関する課題.....	10
3. アクセシビリティに対応した電子書籍コンテンツ制作のあり方 .....	12
3.1. 電子書籍の制作プロセスにおける音声読み上げ対応の実施工程.....	12
3.1.1. 音声読み上げ対応手順.....	12
3.1.2. 音声読み上げ対応の実施場面 .....	12
3.1.3. 読み上げ指定方法 .....	13
3.1.4. 音声読み上げ対応の実施手順 .....	13
3.2. 読み上げ指定及び校正のための制作支援ツール .....	14
3.2.1. 制作支援ツールの機能 .....	14

3.2.2.	各機能に求められる要件 .....	14
3.3.	出版社における音声のチェック .....	15
3.3.1.	チェック手順.....	15
3.3.2.	音声校正用紙面データ .....	16
3.3.3.	音声用の校正記号 .....	17

# 1. 読み上げ対応のための電子書籍記述仕様

※本章では、現在の EPUB3 の仕様と比較し、アクセシビリティに配慮した電子書籍の音声読み上げを、より高品質かつ効率的に実現するために必要となる属性について検討した結果について記載している。これら属性に関して、総務省としては今後標準化への取組について支援していく予定であるが、現段階においてはあくまで検討結果にすぎず、推奨するものではないことにご注意されたい。

## 1.1. 音声読み上げのために必要なメタデータ

音声化情報を付与するために、EPUB3 で規定されている既存の属性に加え、音声読み上げに必要なメタデータを定義する。

既存の音声関連属性には、Alphabet と Phoneme がある。また PLS では、この 2 つに加え、grapheme という属性を持つ。この 3 つの属性情報により、ある文字がどのような発音をするのかを定義することができる。しかし、日本語の場合、英語などの他の言語との併記を行う場合も多く、日本語だけの言語指定ではうまく表現ができない場合がある。このため、SSML の仕様では規定があり、EPUB3 で規定のない属性のうち、言語を制御することのできる属性を定義する。

言語指定 (lang 属性) : 読み上げる言語を指定する

発話者指定 (gender 属性) : 読み上げ音声の声の種別 (男声/女声等) を指定する

### 1.1.1. lang 属性 :

通常の lang 属性は、xml:lang として、その文章の発音する言語 (language) が何であるかを規定するものである。例えば、語学教材であった場合、日本語表記とそれに対応する外国語の表記が必要となる。この規定により、外国語は外国語の発音で、日本語は日本語の発音で表現することができるため語学教材の音声化には必須となる。

TTS ソフト起動時のオーバーヘッドを短縮するため、リーダーが電子書籍を読み込んでいる間に、利用する TTS ソフトを事前に起動して待機させる場合の対応として、EPUB3 のマニフェストファイル (toc ファイル) にあらかじめ利用する TTS ソフトを記載する。これにより、TTS ソフトの起動の準備が可能となる。基本的にマニフェストファイルに記載することを推奨するが、その際、xml:lang で記載してしまうと、その文章が xml:lang で規定された言語ということになってしまい、あらかじめ、複数の言語を指定することができない。そこで、現在の EPUB3 には規定がない、ssml:lang や voice:language 等の新しい属性を設定することで、TTS ソフトの事前起動を可能にする。

### <記述例>

日本語の TTS ソフトと英語の TTS ソフトを利用する場合の一例

```
<metadata ssml:lang="jp" ssml:lang="en" …>
…
</metadata>
```

なお、XHTML ファイルで規定する `xml:lang="jp"` は、その XHTML ファイルの言語を指定するものとして必要であるため、デフォルトで日本語にする場合は、XHTML ファイルの定義部分で以下のように宣言すればよい。`ssml:alphabet` を規定しておくこと、XHTML ファイル内の個別の発音指定の部分で宣言を省略することができる。

<記述例>

```
<html xmlns:ssml="http://www.w3.org/2001/10/synthesis"
      xml:lang="jp"
      ssml:alphabet="x-JEITA">
```

※ `ssml:alphabet` で規定している "x-JEITA" については、次節で解説する。また `xml:lang` で指定する、言語コードは、RFC4646 で規定されている。

#### 1.1.2. gender 属性 :

gender 属性についても、lang 属性と同様に記載する。

### 1.2. 音声読み上げ指定技術の用途と適用方法

文章に音声化情報を指定するための方法として、全文 SSML と部分 SSML を利用する。

以下に、2 つの用途と適応方法について定義する。

#### 1.2.1. 全文 SSML と部分 SSML の用途

##### ➤ 全文 SSML の用途

全文 SSML の用途としては、デジタル教材・教科書、資格試験テキスト、実用書など、文字の読みを間違えて発音されると問題が起きる分野の電子出版物、及び公的機関の発行する白書や行政資料など情報保障の観点で対応を迫られる分野の電子出版物が挙げられる。

なお、全文 SSML は、発音辞書 (PLS) の充実度により部分 SSML で代行可能であるが、一部再生側の TTS ソフトに読み上げが依存する場合もある。

##### ➤ 部分 SSML の用途

部分 SSML の用途としては、文芸、ビジネス書など一般大衆向け出版の分野の電子出版物が挙げられる。固有名詞 (人名、地名)、専門用語、当て字以外では、文字の読み間違いが少ないと想定されるため、発音辞書 (PLS) への登録及び部分 SSML による指定で大部



分解決すると思われる。コストがかかりすぎると出版社としては取り組めないため、部分 SSML が採用される書籍は、カテゴリごとに異なると考えられる。なお、文中には SSML を埋め込まず、発音辞書 (PLS) への単語登録だけに特化するやり方もある。

### 1.2.2. SSML の適用方法

SSML の適用方法には、全文に SSML を適用する場合、部分的に SSML を適用する場合、発音辞書 (PLS) のみを適用する場合の 3 種類がある。以下、それぞれについて解説する。

#### ➤ 全文 SSML

再生側の端末や PC にアプリケーションとしてインストールされている TTS ソフトの環境によって読み方が変わってほしくない場合 (著者や出版社が意図した通りに読む場合) に用いる方法である。全文 SSML の適用方法としては、メディアオーバーレイ (Media Overlays) の実現と、TTS ソフトの再生単位を明確にするために、句単位で SSML を付与することとする。

#### ➤ 部分 SSML

部分 SSML は、発音辞書に登録すると他の位置で間違った読み方をする場合や、発音辞書に登録された語句とは別の読み方をさせたい場合に適用する方法である。

文芸作品で、例えば、宮本武蔵の場合、小説のはじめでは、“たけぞう” と読ませているが、途中から“むさし” と変わる。完全に換わってしまえばいいが、“むさし” と換わった後でも、“たけぞう” と読ませる場合がある。こうした場合は、武蔵を“むさし” で発音辞書に登録し、“たけぞう” と読ませたい場所に部分 SSML で“たけぞう” と適用する。こうすれば、通常は“むさし” と発音され、部分 SSML が適用された場所は、“たけぞう” と発音される。

また、現在流通している TTS ソフトは、非漢字及び、第 1、第 2 水準漢字 (JIS X 208:1997) 以外の漢字コードを文字として扱わない場合がほとんどである。そのため、JIS X 208:1997 以外の漢字コードを利用する場合は、部分 SSML を適用することで正確な読みを実現する。

#### ➤ 発音辞書 (PLS) のみ

固有名詞、地名、専門用語、当て字など出版物ごとに一意に読みが決まる語句については、部分 SSML で対応すると適用した部分のチェックなどでコストがかかるため、出版校正を行う際に、その出版物で出てくる固有名詞 (人名、地名)、専門用語、当て字などの単語を抽出し、発音辞書に登録すれば、語句一つに対して、1 回のチェックで済む。なお、発音辞書に登録する際の留意点は、以下の通りである。

- ✓ 固有名詞（人名、地名）、専門用語、当て字に限定して登録する  
TTS ソフトは通常、大規模なシステム辞書を内蔵している。発音辞書に一般的な用語を登録すると、アクセントの位置ずれや、イントネーションがおかしくなる場合がある。そのため、登録する単語については、上記の通り、固有名詞（人名、地名）、専門用語、当て字に限定して登録することが望ましい。ただし、同じ表記で別の発音をする固有名詞を登録する際は注意が必要である。例えば、兵庫県三田（さんだ）市にある三田（さんだ）市立図書館を、略して、三田図書館と書くと、東京都港区にある三田（みた）図書館と区別できなくなるため、発音辞書に登録するかどうかを吟味する必要がある。
  
- ✓ 句読点を含む文章は発音辞書に登録しない  
TTS ソフトは、入力されたテキストを解析するために、句読点などの読み制御情報を活用して処理している。単語ではなく、句読点を含む文章を登録すると、このテキスト解析を円滑に行う障害要因になる場合がある。また点字変換する際にも、発音辞書が利用されることから、発音辞書には、句読点を含まない単語単位の登録を行うものとする。テキストから点字に点訳する際も同様に、TTS ソフトと同じように文書解析を行い、分かち書きで表現するため、基本的に単語単位で区切る必要がある。
  
- ✓ 登録する単語の表記は長い方が望ましい  
TTS ソフトは、発音辞書を優先して発音しようとするため、あまり短い単語を登録すると、入力テキストの解析制度を下げる危険があり、意図しない読み方をする場合がある。例えば、米大統領の米の部分を“べい”ではなく、“アメリカ”と読ませたい場合、「米大統領」＝“アメリカだいてうりょう”と登録すれば問題はないが、「米」＝“アメリカ”と登録すると、米と書いてあるところはすべて、“アメリカ”と発音されてしまう。

## 2. アクセシビリティ対応電子書籍リーダー設計指針

SSML を付記した EPUB を解釈するための電子書籍リーダー（以下、リーダーと略す）の設計指針を示す。

### 2.1. SSML 及び PLS の表記方法

#### 2.1.1. 音声表記について

音声発音の記述言語は現在、IPA、JEITA、SAMPA、L&H+が利用されているが、本指針では、日本語については、JEITA を利用することとし、外国語は IPA を利用する。

JEITA では、「日本語テキスト音声合成用記号」の規格として、日本語の取り扱いを便利にするための「仮名レベルの表記」を採用している。日本の出版物のほとんどが日本語の取り扱いのため、この仮名レベルの表記を利用するために、JEITA を採用することが適切である。また JEITA では、仮名レベルよりも詳細な表記が可能な「異音レベルの表記」も採用している。この表記は、国際音声記号（International Phonetic Alphabet: IPA と略す）を基本としているため、言語指定には、IPA を利用することが最も適している。外国語の場合の表記は、IPA とする。

日本語の場合： `<span ssml:alphabet="x-JEITA" ssml:ph="*****"> **** </span>`

※ ph は、Phoneme を省略したもの

英語の場合： `<span ssml:alphabet="IPA" xml:lang="en"> **** </span>`

#### 2.1.2. SSML の表記方法

全文 SSML 及び部分 SSML の文字エンコーディングにおいて、半角カナの利用は UNICODE 変換の際に問題となるため、利用を禁止し、全角カナで記述する。なお文字エンコーディングは、UTF-8 を推奨とする。

#### ✓ 全文 SSML の場合

全文 SSML では、句（句点まで）を一つの単位として扱う。句単位で、span タグを利用して、発音情報を埋め込む。

<記述例>

```
<p>
<span ssml:ph="チューゴクギョセンガ/オガサワラ/キンカイニ/シュ>ツボツ>シ>ハジメタノハ
コンシュンカラデス.>">中国漁船が小笠原近海に出没し始めたのは今春からです。</span>
<span ssml:ph="クガツ/カラ/キューゾーシ|...">9月から急増し、・・・</span>
</p>
```

✓ 部分 SSML の場合

部分 SSML では、単語を一つの単位として扱う。単語単位で、span タグを利用して、発音情報を埋め込む。

<記述例>

```
<p>
<span ssml:ph="コノエ/フミマロ">近衛文麿</span>は、
<span ssml:ph="ヒガシ/クニノミヤ">東久邇宮</span>内閣にて、国務大臣として入閣した。
大日本帝国憲法改正に意欲を見せたものの、A 級戦犯に指定され服毒自殺した。
</p>
```

人名等の固有名詞は、PLS で規定する方が制作及び校正に手間がかからないと考えられる。

### 2.1.3. PLS の表記方法

PLS ファイルのフォーマットは、W3C の PLS 仕様に記載の通り、以下のように記載する。

<記述例>

```
<lexicon
  version="1.0"
  alphabet="x-JEITA"
  xml:lang="jp"
  xmlns="http://www.w3.org/2005/01/pronunciation-lexicon">
  <lexeme>
    <grapheme>近衛文麿</grapheme>
    <phoneme>コノエ/フミマロ</phoneme>
  </lexeme>
  <lexeme>
    <grapheme>東久邇宮</grapheme>
    <phoneme>ヒガシ/クニノミヤ</phoneme>
  </lexeme>
</lexicon>
```

上記のデータは、発音辞書のデータファイルとして、制作支援ツールで作成するため、上記のファイルを直接管理することは避けられるが、作業が慣れてくると直接ファイル进行操作したくなる場合がある。そのような場合は、簡易的に、以下のような<phoneme>タグによって、データを管理し、自動変換ツール等によって、上記の PLS フォーマットに変換すればよいと考える。あくまで支援ツールの拡張機能として実現すればよい。

#### <記述例>

```
<phoneme alphabet="x-JEITA" ph="コノエ/フミマロ">近衛文麿</phoneme>  
<phoneme alphabet="x-JEITA" ph="ヒガシ/クニノミヤ">東久邇宮</phoneme>
```

PLS データの記載における留意点は以下の通り。

- ✓ PLS の文字エンコーディングにおいて、半角カナの利用は UNICODE 変換の際に問題となるため、利用を禁止し、全角カナで記述する。なお文字エンコーディングは、UTF-8 を推奨とする。
- ✓ 登録語句は、点字変換を考慮して、単語単位とする。

#### 2.1.4. 間の制御方法

SSML 表記における間を制御は、JEITA 規格 (JEITA IT-4006 の日本語テキスト音声合成用記号) の記載に準じる。ここでは、SSML を利用して、文中、文間の息継ぎ、あるいは話し方の間 (ま) などによって生じる音声の休止区間の表現について記載する。例えば、「1、はじめに」と1とはじめにの間に一呼吸おいて、発音させたい場合は、“\_”もしくは“\_ 1”をポーズとして挿入する。

#### <記述例>

```
<span ssml:ph="イチ_ 1 ハジメニ. "> 1 はじめに</span>  
<span ssml:ph="_ 1 "></span>
```

#### 2.1.5. SSML 及び PLS の本文への埋め込み方法

SSML 及び PLS における発音記号のテキストを埋め込み作業は、手動で行うことも可能である。しかし、作業の効率性と品質を高めるために、制作支援ツール (音声化情報の付与支援ツール) を整備して、自動もしくは半自動で行った方がよい。

#### 2.2. リーダー側に求められる機能

SSML を付記した EPUB を音声再生するため、リーダーが動いている装置 (タブレット、専用端末、PC など) に、TTS ソフトを組み込む必要がある。ここでは、リーダーと TTS ソフトとの間でどのようなやり取りをするかを規定する。

##### 2.2.1. リーダーの基本機能

EPUB をリーダーが読み込むと、まず初めに EPUB の構造を解析する。EPUB に DRM がかかっている場合は、通常リーダー側が初期処理として内部メモリ空間に DRM 解除の状態 (暗号のかかっていない EPUB) を作成する (描画や構造解析のために必要な処理であり、アクセシビリティに特化した処理ではない)。リーダーから TTS ソフトへ SSML 付きテキストをある単位 (文節単位、句

単位、段落単位など)で渡す。TTS ソフトはリーダーから渡された SSML 付きの文章を解釈し、音声化を行う。

### 2.2.2. 音声データと文字との同期機能

音声ファイル付き EPUB の場合は、音声データと文字とを同期させるために、SMIL を利用しているが、SSML で記載した場合は、あらかじめ発声時間と文字との同期をとれないため、SMIL で連動させることができない。そのため、リーダーと TTS ソフトとの間で SSML 付きのデータのやり取りを規定する。

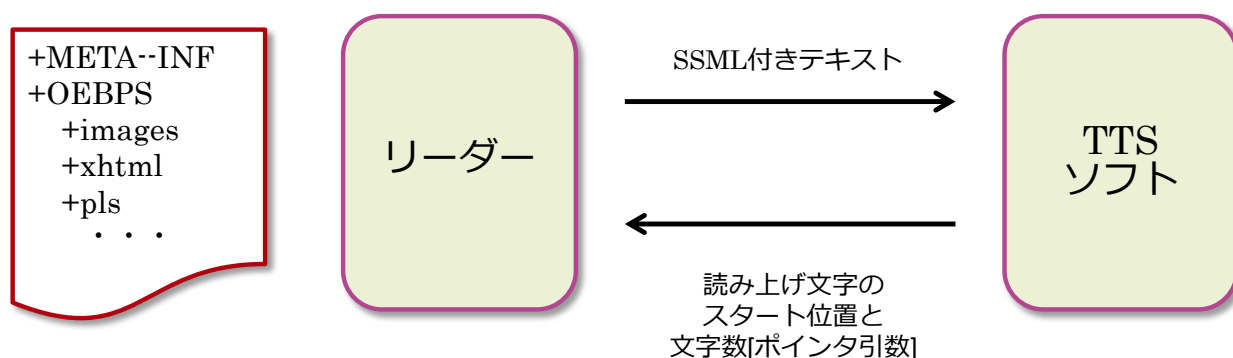


図 1 リーダーと TTS ソフトとの SSML 付きテキストの受け渡しと返り値

図 1 に示すように、リーダーは、SSML 付きテキストを TTS ソフトに渡し、その結果読み上げ文字のスタート位置と文字数をポインタ引数として取得する。これによってリーダーは実際どこが読み上げられているかを認識することが可能となる。

### 2.2.3. 複数の本を同時に開く機能

リーダーが、複数の EPUB を同時に開く場合もあるため、リーダーに接続する TTS ソフトの実装としては、マルチスレッド方式を採用し、複数の本が同時に開かれた場合にも、個別のスレッドとして TTS ソフトを稼働することができるようにすることが望ましい。これは発音辞書における登録語句の読み上げ矛盾を避けるためである。

### 2.2.4. 発音辞書ファイル (PLS) の登録及び解除機能

発音辞書ファイル (PLS) は、リーダーが TTS ソフトと初期同期を行ったとき、一番初めに渡されるデータファイルである。リーダーは、TTS ソフトが用意している発音辞書登録機能 (API) を呼び出し、PLS を TTS ソフトに渡す。TTS ソフトは、PLS を受け取ると初期化の時に内部辞書にマージし、音声化の読みデータとする。図 2 に発音辞書の登録イメージを示す。

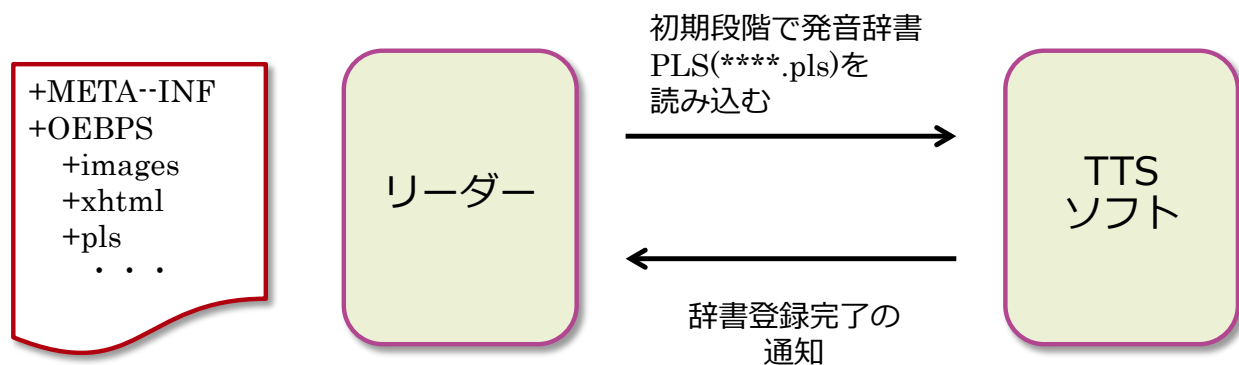


図 2 リーダーと TTS ソフトとの発音辞書(PLS)の受け渡しと返り値

一つの EPUB に、本来であれば複数の発音辞書ファイルを持つことが可能であるが、当初は、リーダーと TTS ソフトとのやり取りを簡素化するために、一つの EPUB に紐づけられる PLS は高々 1 つとする。

また各 XHTML ファイル（EPUB における各ページ）内で、リンクタグ(link)によって PLS を指定しなくても、すべての XHTML ファイルで有効になるものとする。

発音辞書の有効期間は、リーダーが該当する EPUB を閉じるまでとする。閉じる場合は、リーダー側から、TTS ソフトの発音辞書登録解除機能（API）を呼び出し、PLS を解除する。

#### 2.2.5. SSML データの受け渡し機能

XHTML ファイルに記載されている SSML 付きテキストをリーダーは、TTS ソフトに渡して、音声化を行う際の留意点を列挙する。

- ✓ テキスト中の<> & ‘ “などの文字は、実体参照記法の形式（&lt; &gt; &amp; &apos; &quot; など）であればディスプレイに表示上問題はないが、TTS ソフトに渡す場合は、実体参照記法で渡すと読めない TTS ソフトが存在する。そのため、実体参照記法の形式で記載されている場合は、TTS ソフトに引き渡す前に<> & ‘ “に変換する必要がある。

- ✓ ルビ（ruby）と SSML の混在の場合は、TTS ソフト側に SSML 部分だけを渡すこととする。

```
<span ssml:ph="アジサイ">
<ruby><rb>紫陽花</rb><rt>あじさい</rt></ruby>
</span>
```

上記の記載の場合、TTS ソフト側に渡す文字列は、以下のようになる。

```
<span ssml:ph="アジサイ">紫陽花</span>
```

- ✓ 語学学習教材などでは、外国語の発音で表現したい場合、ある文章の読み上げを日本語から外国語に変えるときには言語だけを指定する。以下の例は、TTS ソフトを日本語モードから英語モードに変える場合を示す。ただし、再生側に英語再生用の TTS ソフトが組み込まれていること

が前提となる。もし英語再生用の TTS ソフトが無い場合は、日本語読みで再生するか、もしくは何も再生しない。

```
<p>
  質問があれば何でも聞いてください。
  <span xml:lang="en">
    Please feel free to ask me any questions.
  </span>
</p>
```

なお、今後 `ssml:lang` が EPUB で規定された場合は、`xml:lang` の部分を `ssml:lang` で代用することも可能である。

```
<p>
  質問があれば何でも聞いてください。
  <span ssml:lang="en">
    Please feel free to ask me any questions.
  </span>
</p>
```

#### 2.2.6. PLS に関する課題

第 3、第 4 水準の漢字及び外字の発音指定について、TTS ソフトでは、現在文字として認識していないため、リーダー側で変換するか、部分 SSML として音声化を行うかが課題となる。

例えば、PLS で規定する場合は、

```
<phoneme alphabet="x-JEITA" ph="ミヤザキ">宮崎</phoneme>
```

と記載すればいいのだが、TTS ソフトが、崎を認識できないため、読みとして変換できない可能性がある。

```
<phoneme alphabet="x-JEITA" ph="ミヤザキ">宮 = </phoneme>
```

データを読み込んだ際、崎の字が、「ゲタ」になる場合がある。これを避けるために、UNICODE 表記を推奨する。

UNICODE で記載があれば、フォントがない外字や異体字の場合を除き、TTS ソフトに組み込む前に漢字変換を行い、音声化処理を行うことが可能となる。しかし、EPUB データ上、画像ファイルが添付してある可能性もあり、第 3、第 4 及び外字の漢字については、現段階では保留する。

第 1 及び第 2 とそれ以外の文字体系を図 3 に示す。



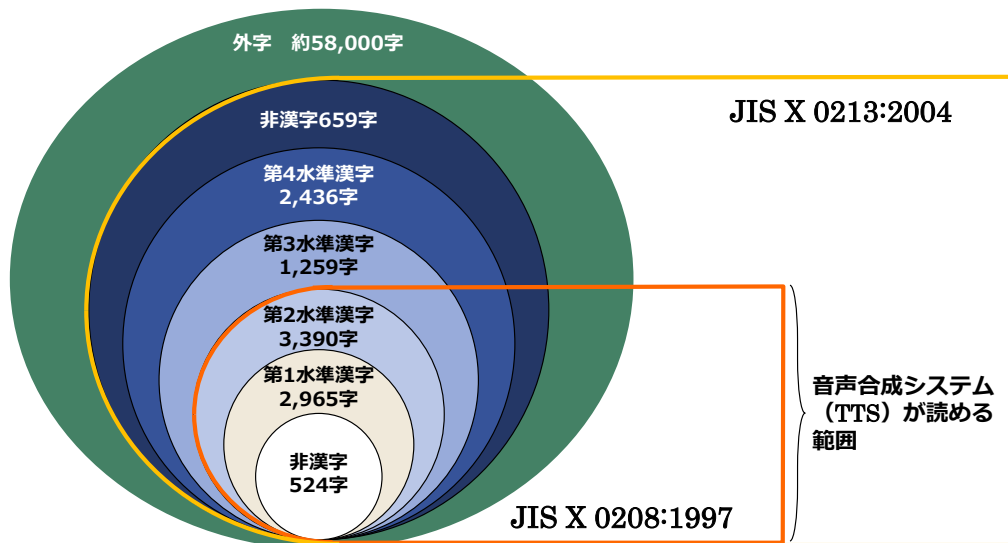


図 3 TTS ソフトの読み上げ可能な領域と読み上げ不可の領域

日本語の文字は JIS 第 1 水準、第 2 水準を規定した JIS X 0208:1997 と、第 3 水準、第 4 水準を規定した JIS X 0213:2004、そしてこれらに含まれない外字などが存在する。TTS ソフトで読み上げ可能な文字は現状では JIS X 0208:1997 の範囲にとどまっており、JIS 化されている文字のほぼ半分が読み上げ対象外となっている。

この課題が解決するまでは、EPUB 本体の中で部分 SSML として記載することで対応可能である。

[EPUB 実装]

今日のゲストは、宮崎あおいさんです。

[EPUB 表示]

今日のゲストは、宮崎あおいさんです。

[音声化]

キョウノ ゲストワ ミヤザキ アオイサンデス。

### 3. アクセシビリティに対応した電子書籍コンテンツ制作のあり方

#### 3.1. 電子書籍の制作プロセスにおける音声読み上げ対応の実施工程

##### 3.1.1. 音声読み上げ対応手順

音声読み上げ対応のため、以下の業務を追加する。

- 1：構文解析ソフト等を用いた、テキストデータの音声情報化
- 2：音声情報化された内容のチェックと読みの修正
- 3：正しい読みの音声情報の電子書籍データへの埋め込み

##### 3.1.2. 音声読み上げ対応の実施場面

音声化については、以下の2通りとする。

- 1：オーサリングの中で、タグ付け作業等と合わせて実施する
- 2：オーサリング後、制作されたファイルに音声化情報を付加する方法とする

電子書籍のオーサリング中に、オーサリング作業の一環として実施するか、電子書籍のオーサリング完了後に別の作業として実施するかは、制作会社の業務の進め方が各社によって異なる。

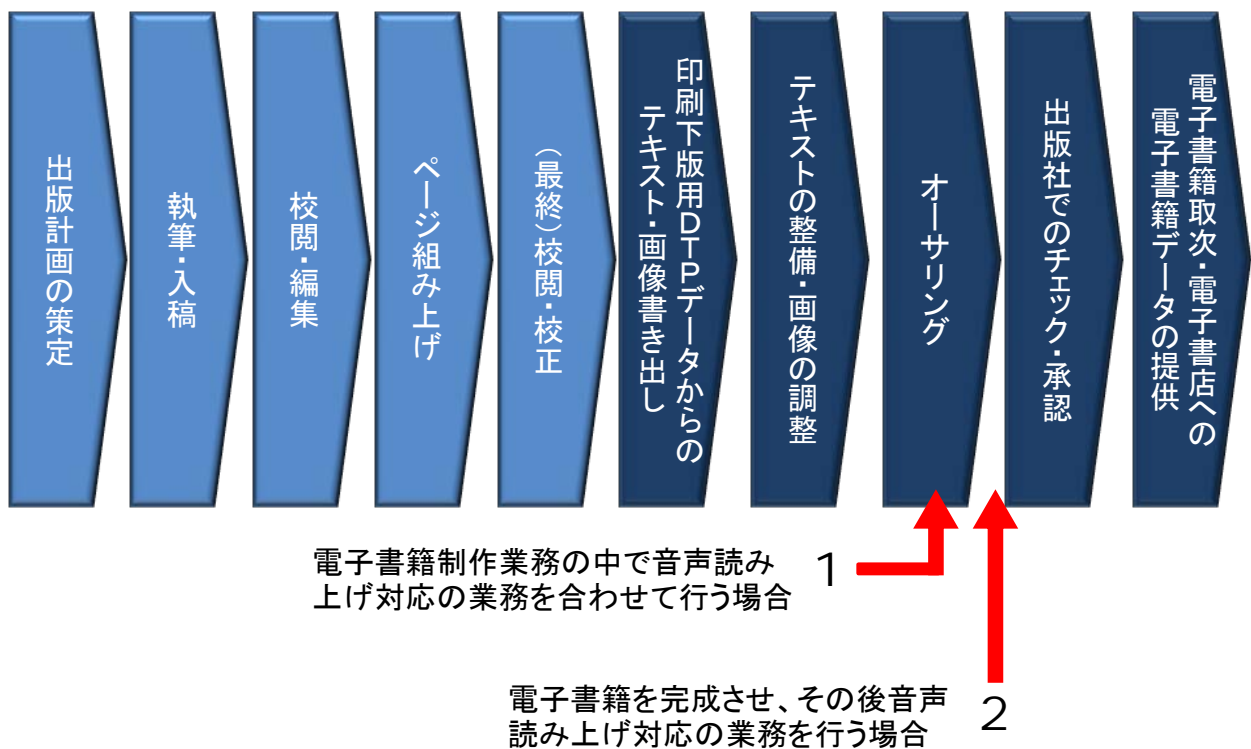


図4 音声読み上げ対応の実施場面

### 3.1.3. 読み上げ指定方法

読み上げ指定方法については、文章全体への読み指定（全文 SSML）、または読み間違いや読み飛ばしが起こりそうな部位のみを対象とした読み指定（部分 SSML）のいずれかとする。

### 3.1.4. 音声読み上げ対応の実施手順

想定される音声読み上げ指定の実施手順を示す。

これらの手順においては、制作支援ツールによる自動処理と、支援ツールを利用した作業支援が必要となる。

- （手順 1）電子書籍用 EPUB ファイルからのテキストデータ取り出し
- （手順 2）テキストデータからの読み上げデータ生成（手順 1、2 は自動処理）
- （手順 3）読み上げデータの確認・修正（支援ツールでの作業）
- （手順 4）出版社でのチェック・承認
- （手順 5）チェック結果の反映（支援ツールでの作業）
- （手順 6）読み上げ用 SSML タグ／PLS の EPUB ファイルへの埋め込み（自動処理）

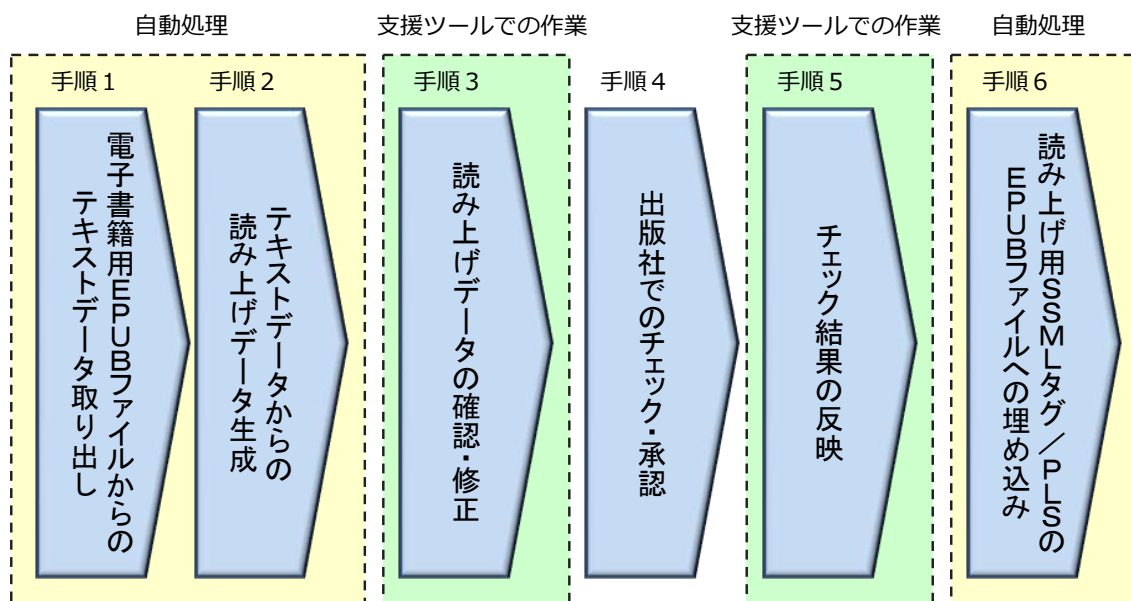


図 5 音声読み上げ対応手順

## 3.2. 読み上げ指定及び校正のための制作支援ツール

読み上げ指定を行うため、制作支援ツールを用いることとする。

制作支援ツールは、以下の機能を持つ。

### 3.2.1. 制作支援ツールの機能

#### (機能 1) 読み上げデータ生成

EPUB データ、またはテキストデータを入力すると、自動で構文解析と読み上げ付与を行い、その結果を出力する機能（手順 1、2 を実施する機能）。

#### (機能 2) 読み上げデータ入力

EPUB データ、またはテキストデータを画面表示し、作業者がキーボードから読みを直接入力する機能（手順 3、5 で使用する機能）。

#### (機能 3) 読み上げデータ修正

機能 1 で出力された読み上げデータや、機能 2 で入力された読み上げデータを画面表示し、作業者がキーボードから修正する機能（手順 3、5 で使用する機能）。

#### (機能 4) SSML 埋め込み

修正が完了した読み上げデータを、EPUB データに SSML タグとして自動で埋め込む機能（手順 6 で使用する機能）。

#### (機能 5) PLS（読み上げ用の変換辞書）生成

読み上げ用辞書を作成し、EPUB データに埋め込むための PLS ファイルを生成する機能（手順 6 で使用する機能）。

### 3.2.2. 各機能に求められる要件

(要件 1) 校正が必要な個所が明示され、簡単に修正が行えること

(要件 2) 技術的な素養（HTML タグや SSML に関する知見等）がなくても、音声読み上げの修正を行うことができること

(要件 3) 音声でなく、紙に印刷して目視でチェックできる仕組みを持つこと

### 3.3. 出版社における音声のチェック

#### 3.3.1. チェック手順

制作会社が読み上げデータを作成し、内部での確認と修正を行ったら、その結果を出版社の出版社側で確認と修正指示を行い、制作会社に返送する。制作会社は修正指示を反映し、出版社側に最終的な確認と承認をしてもらう。

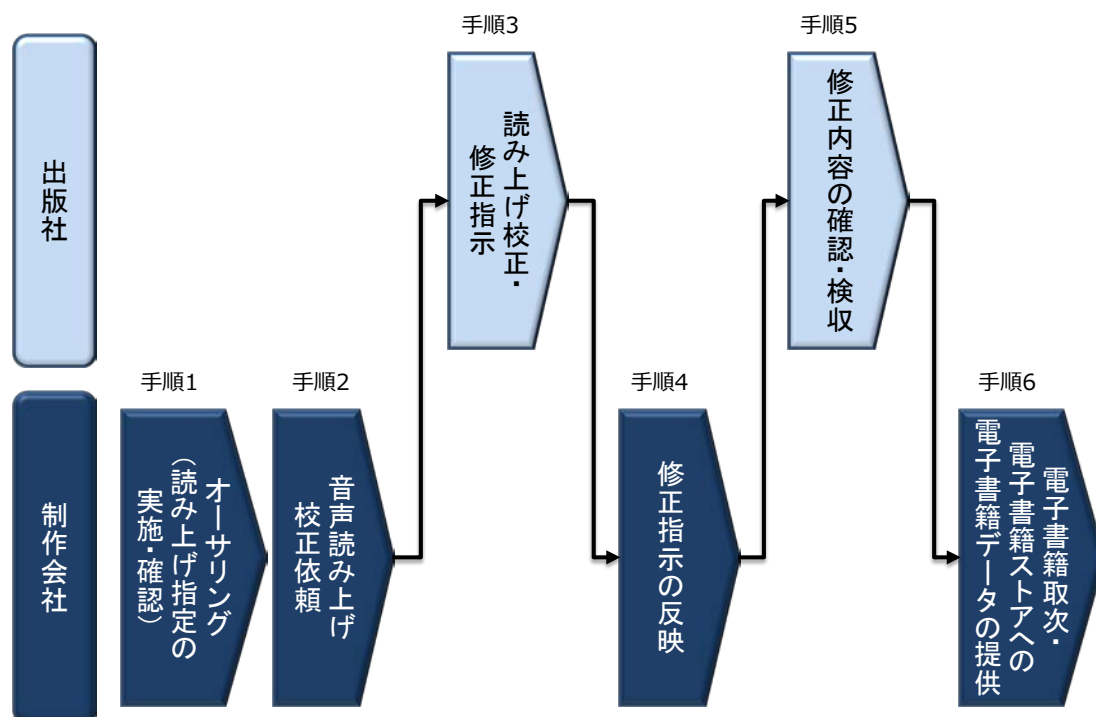


図 6 音声版電子書籍のチェック手順

図 6 の解説：下記のチェック手順 1～6 のフローを図示。

手順 1：オーサリング（読み上げ指定の実施・確認）【制作会社】

手順 2：音声読み上げ校正依頼【制作会社】

手順 3：読み上げ校正・修正指示【出版社（版元）】

手順 4：修正指示の反映【制作会社】

手順 5：修正内容の確認・検収【出版社（版元）】

手順 6：電子書籍取次・電子書店への電子書籍データの提供【制作会社】

### 3.3.2. 音声校正用紙面データ

出版社と制作会社のやり取りは音声データではなく、紙面上での発音表記の確認によって行う。

電子書籍の紙面に直接読み上げ情報が記入されると、紙面が煩雑になりチェックすべき箇所が分かりにくいほか、校正指示を記入する余白等が十分に取れない恐れがあることから、紙面データとは別に読み上げデータを印刷できる形式で提供することも認める。

紙面データ	読み上げデータ
<p>① 二次方程式とは、次数<sup>②</sup>2の代数方程式のことである。定義に従えば、</p> $ax^2 + bx + c = 0$ <p>(<math>a \neq 0, b, c</math> は定数)と表される。これを二次方程式の一般形 (generalized form) ④という。さらに二次方程式について、いくつかの特徴をもつ特殊な形が考えられる。⑤本項では便宜的に以下の用語を用いる。</p>	<p>① ニジホウテイシキトワ、ジスウニノ/ダイスウホウテイシキノコトデアル。</p> <p>② テイギニ/シタガエバ、エーエックスニジョー/タス/ビーエックス/タス/シー/イコール/ゼロ_カッコ/エー_ノットイコール/ゼロ/テン_ビー/テン_シー/ワ/テースー/ト_アラワサレル。</p> <p>③ コレオ/ニジホーテーシキノ/イッパンケー/カッコ/ジェネラライズド/フォーム/カッコジ/トイウ。</p> <p>④ サラニ/ニジホウテーシキニ/ツイテ、イクツカノ/トクチョーオ/モツ/トクシュナ/カタチカ/カンガエラレル。</p> <p>⑤ ホンコーデワ/ベンギテキニ/イカノヨーゴ/オ/モチール。</p>

図7 音声校正用の紙面データ例

図7の解説：紙面データの印刷イメージと読み上げデータの印刷イメージを図示。

修正指示は版元の校正担当者が紙面上に赤字等で記入し、制作会社に送付する。制作会社側では、この指示に基づき、端末上で読み上げデータの修正を行う。

<p style="text-align: center; color: red;">アケル</p> <p>① <del>ニジホウ</del>テイシキトワ、ジスウニノ/ダイスウホウテイシキノコトデアル。</p> <p>② テイギニ/シタガエバ、エーエックスニジョー/タス/ビーエックス/タス/シー/イコール/ゼロ_カッコ/エー_ノットイコール/ゼロ/テン_ビー/テン_シー/ワ/テースー/ト_アラワサレル。</p> <p>③ コレオ/ニジホーテーシキノ/イッパンケー/カッコ/ジェネラライズド/フォーム/カッコジ/トイウ。</p> <p>④ サラニ/ニジホウテーシキニ/ツイテ、イクツカノ/トクチョーオ/モツ/トクシュナ/カタチカ/カンガエラレル。</p> <p>⑤ ホンコーデワ/ベンギテキニ/イカノヨーゴ/オ/モチール。</p>
--

図8 音声の修正指示

図8の解説：図7の読み上げデータの印刷イメージに校正赤字を入れた状態を図示。

紙面と読み上げ情報が別のシートの場合は、出版社側での校正がやりにくいことも考えられる。その場合は読み上げ校正用シートに紙面情報と読み上げ情報を一緒に記載する方法も認める。

ただし読み上げ校正用シートでは紙面レイアウトが崩れてしまうことから、別途紙面レイアウトを確認するためのシートも送付する。

### 入力紙面確認シート

二次方程式とは、次数2の代数方程式のことである。定義に従えば、  
 $ax^2 + bx + c = 0$   
 ( $a \neq 0, b, c$ は定数)と表される。これを二次方程式の一般形 (generalized form) という。さらに二次方程式について、いくつかの特徴をもつ特殊な形が考えられる。本項では便宜的に以下の用語を用いる。

### 読み上げ校正用シート

ニジホーテーシキトワ、ジスーニノダイスーホーテーシキノコト  
 二次方程式とは、次数2の代数方程式のこと  
 デアル。テーギニシタガエバ、  
 である。定義に従えば、  
 エーエックスノジジョーノプラスノピーエックスノプラスノシーノ  
 $ax^2 + bx + c = 0$   
 イコールノゼロ  
 カッコノエーノットイコールゼロノテン\_ピーノテン\_シーノワノ  
 (  $a \neq 0$  ,  $b$  ,  $c$  は  
 テースト\_アラワサレル。コレオノニジホーテーシキノ  
 定数) と表される。これを二次方程式の  
 イッパンケーノカッコノジェネライズド\_フォームノカクコジ  
 一般形 ( generalized form )

図9 音声校正用の紙面データ (別パターン)

図9の解説： 入力紙面に総ルビの形で読み上げ情報が付いた校正シートのイメージを図示。

### 3.3.3. 音声用の校正記号

音声用の校正記号の定義：

出版社と制作会社間で、音声表記を印刷出力した上で音声校正を行う場合、分かりやすく、簡単に修正指示が行えるよう、紙面上に記入するための音声用の校正記号を定義する。

読みの修正： 訂正箇所を棒線で消し、正しい読みをひらがなまたはカタカナで記入

ポーズの挿入： ポーズを示す記号（「/」、「\_」等）や、文字で指示を記入（「アケル」等）

音声合成用表記記号では厳密な記述仕様が定まっているが、読み記号、韻律記号を電子書籍の校正担当者が理解し、使用することについては極めて障壁が高いため、紙面の校正と共通性の高いルールが適切である。