

TEI研究会2023-09 研究会の記録

開催日時

2023年6月1日(木)18-20時

参加者

氏名(所属等)

永崎 研宣 (一般財団法人人文情報学研究所)

中川 奈津子(国立国語研究所)

曹 芳慧 (大阪大学人文学研究科博士後期課程)

李 媛 (関西大学KU-ORCAS)

纓田 宗紀 (アーヘン工科大学)

王 雯璐 (東京大学東京カレッジ)

藤原 静香 (京都女子大学大学院／京都産業大学非常勤研究員)

宮川 創 (国立国語研究所)

王 一凡 (東京大学教育学研究科／人文情報学研究所)

塚越 柚季 (東京大学)

山元 啓史 (東京工業大学)

石田 友梨 (岡山大学学術研究院社会文化科学学域)

主な話題

1.TEIに関する報告

①纓田

「西洋中世学会」TEI入門セミナーを開催予定

2.『方言談話資料』マークアップ

【概要】

- 『[方言談話資料](#)』データ | [国立国語研究所](#)、「京都」から作業を開始する。
(「国立国語研究所資料集 10 方言談話資料 (4) 一福井・京都・島根」1981年1月)
- 「[oxygen](#)」の画面を共有しながら、マークアップについて検討を行った。
- 将来的には、音声ファイルをid単位で区切って対応させる予定でいる。

【参考】[2023-05議事録](#)／[2023-07議事録](#)／コプト語コーパス[besa.letters_TEI](#)

【マークアップの実践】

- まずは、[2023-07議事録](#)を見ながらタグ付けの再確認を行った。
- phrタグすべてに”type=文節”を付けるのは見た目が悪い。ではどうするか？
⇒tagUsageを利用して、そこに説明を示すという手段を検討した。

(いかに入力する回数を減らせるかが大事＝入力ミスも減る)

- tagsDeclについて
[TEI element tagsDecl \(tagging declaration\)](#)
- tagUsage
@gi: 当該タグで示された要素の名前

```
<encodingDesc>
  <tagsDecl>
    <namespace name="http://www.tei-c.org/ns/1.0">
      <tagUsage gi="phr">
        <gi>phr</gi> タグは文節区切りとして用いる。
      </tagUsage>
    </namespace>
  </tagsDecl>
</encodingDesc>
```

- [歴史的地名／現代地名による境界データ検索 | Geoshapeリポジトリ](#)
⇒地名情報を入力するときに利用できるかもしれない。

- 凡例の記述方法について検討を行った。

①注記

＝noteタグを使用する

〈候補Ⅰ〉

```
<note type="footnote" n="1">Cの発言、重複して聞かれず。</note>
```

⇒文末注であることを示すことにしたため、次のように記述を変更した

〈候補Ⅱ:採用〉

```
<note n="1" place="end">Cの発言、重複して聞かれず。</note>
```

②聞き取り困難な場合

＝TEIガイドライン([8 Transcriptions of Speech - The TEI Guidelines](#))から該当するタグを検討する

〈候補Ⅰ:採用〉

unclear([TEI element unclear \(unclear\)](#))。以下は[2023-05議事録](#))

※<unclear>エレメントの@reason属性には推奨される値が提供されている。たとえばOxygen XML Editorでこの属性を入力しようとすると以下の例のように値の一覧がリスト表示され個々の説明も表示される。

いへる<unclear reason="faded">心ば</unclear>へして、<ruby>
 <rb>朽葉色</rb>
 <rt>くちばいろ</rt>
</ruby>の<ruby>
 <rb>狩衣</rb>
 <rt>かりぎぬ</rt>
</ruby>の袖を台にて、<ruby>
 <rb>烏帽子</rb>
 <rt>えぼし</rt>
</ruby>

background_noise	(eccentric_ductus) indicates illegibility due to an unusual, awkward, or incompetent execution of a glyph or glyphs
eccentric_ductus	
faded	
illegible	
inaudible	

(Fig.1 Oxygen XML Editorにて@reason属性の値の一覧が表示される例)

〈候補Ⅱ>gapタグを使用。[TEI element gap \(gap\)](#)

③笑い・咳

= [8 Transcriptions of Speech - The TEI Guidelines](#)を参照した

```
<u xml:id="u0416913" who="#E">  
  〈スクナイカラ〉。(A エー。) チョード イー トコロエ。ソーユー トコロカ° イー。  
  <vocal>  
    <desc>《笑》</desc>  
  </vocal>  
</u>
```

⇒原文に忠実に《笑》と記述したが、“laugh”と記述したほうが、日本語圏外でも利活用されやすいかもしれない

⇒今後も検討が続く可能性を考慮し、本記述についてはencodingDeclに注記を行うことにした(後々、置換することも可能)。

④(突然の訪問者・電話などで)録音が中断した場合

=<提案 I >gap [TEI element gap \(gap\)](#) /ヘッダーに一括で記述できる

<提案 II : 採用>incident

<https://tei-c.org/release/doc/tei-p5-doc/ja/html/examples-incident.html> /その場を状況を逐一記述できる

【質疑】

- 文節は<phr>ではなく<w>ではないか？
＞文節を「word」については解釈の相違が大きいため、現時点では保留してphrを使用し、随時ヘッダーで説明を加える方針とする

▼次回以降、凡例に基づいて参加者で分担して記述し、とりまとめる作業に入る