Welcome > Editorial Policies > Tagging Guidelines > Foreign Words

Foreign Words

Encoding the Primary Language of the Text

The primary language of the document is indicated by adding a `lang` attribute to the `<body>` element. It is mandatory in *all* transcriptions.

Encoding the Other Languages in the Text

`<foreign>` is used to encode text written in languages other than document's primary language.

- `lang` is used to indicate the language of the foreign text

Defining Languages in the Header

The value of the lang attributes must first be defined (as shown below) within within the header of the document before they can be used within the text.

```
<profileDesc>

  <langUsage>

    <language id="en">English</language>

    <language id="langAbbr">Language Name</language>

  </langUsage>
```

`<profileDesc>` is an immediate child of `<teiHeader>`.

The declaration of language information is required for each language in the text (including the primary language). The abbreviations used as the id values should (wherever possible) be based on the ISO 639-3 three letter abbreviations. If the language in the does not appear within this list, a new abbreviation for that language will be created and added to the tagging and transcription guidelines.

Find out about encoding Special Characters