

EMOXG – Syntactic Hints

Paolo Baggia
Enrico Zovato

Overview

☐ Syntactic Hints from other specifications:

- ☐ EMMA: attribute values related to “xsd” types
- ☐ PLS: Extensibility of attribute values

Definition of Attribute Values

In EMMA 1.0 specification all attribute values are defined as datatypes, examples are:

- `id` of type `xsd:ID` and `ref` of type `xsd:anyURI` (Sect. 4.1.1)
- `composite` of type `xsd:boolean` (Sect. 4.1.2)
- `emma:tokens` of type `xsd:string` (Sect. 4.2.1)
- `emma:lang` of type `xsd:language` (Sect. 4.2.5)
- `emma:confidence` of type `xsd:decimal` in range 0.0 to 1.0 (Sect. 4.2.8)
- `emma:start` of type `xsd:nonNegativeInteger` (Sect. 4.2.10.1)

And also for white space separated multiple values:

- `emma:medium` and `emma:mode` of type `xsd:nmtokens` (Sect. 4.2.11)

■ **EMMA 1.0 Specification**
<http://www.w3.org/TR/emma/>

Extensible Attribute Values (QName)

How to extend the set of values of an attribute?

Solution adopted in PLS 1.0 for “role” attribute of “lexeme”:

- Use of QNames as defined in Section 4 of Namespaces in XML
- “A QName in the attribute content of the “role” attribute is expanded into an expanded-name using the namespace declarations in scope for the containing <lexeme> element. Thus, each QName provides a reference to a specific item in the designated namespace.”

Example:

```
<lexicon version="1.0" alphabet="ipa" xml:lang="en"
  xmlns="http://www.w3.org/2005/01/pronunciation-lexicon"
  xmlns:claws="http://www.example.com/claws7tags">
  <lexeme role="claws:VVI claws:VV0 claws:NN1">
    <grapheme>read</grapheme>
    <phoneme>ri&#x02D0;d<!-- same as ri□d --></phoneme>
  </lexeme>
  <lexeme role="claws:VVN claws:VVD">
    <grapheme>read</grapheme>
    <phoneme>red</phoneme>
  </lexeme>
</lexicon>
```

■ **PLS 1.0 Recommendation**
<http://www.w3.org/TR/pronunciation-lexicon/>

QName - Considerations

- An attribute can define unqualified values
 - ➔ Basic set of values
- An attribute can be extended with different set of values
 - ➔ Use of QName
- QName is defined in a namespace
 - ➔ It is unique, no clashes