Language Documentation and Standards in Digital Humanities: TEI and the documentation of Mixtepec-Mixtec

Jack Bowers, Laurent Romary

Language Documentation and Standards in Digital Humanities: TEI and the documentation of Mixtepec-Mixtec

Jack Bowers
jack.bowers@oeaw.ac.at
Austrian Center for Digital Humanities (ACDH)
Inria
Laurent Romary
Inria

INVENTEURS DU MONDE NUMÉRIQUE

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Mixtepec-Mixtec (Sa'an Savi)

- Sa'an Savi '*rain language*'
- ISO 639-3 code: 'mix'
- Oto-Manguean, Mixtecan, Mixtec-Cuicatec, Mixtepec-Mixtec

• San Juan de Mixtepec Juxtlahuaca district (Oaxaca, MEX)

• Spoken data mostly collected in sessions working with speakers from a small village called Yucunani in the San Juan Mixtepec municipality

• Estimated (+-7,600 speakers)

Source: INEGI (2010)

Has been studied by:

- Pike and Ibach (1978); Paster and Azcona (2004-2007); Beckman and Nieves-SIL (2005-current)

Desired Outcomes

- Create an open source body of reusable and extensible collection of multimedia language resources in the Mixtepec-Mixtec language
- Further the knowledge of all aspects of the language itself
- Demonstrate and evaluate the application of encoding and description standards on a rich but complex collection of lexical and knowledge resources on an under-resourced non-Indo-European language
- Produce and publish empirical corpus-based descriptions and analyses of various aspects of the language's features
- Demonstrate and test the application and utility of descriptive features from cognitive linguistics such as those used to describe Mixtec in the literature in the annotation of the corpus

Basic Challenges in Studying Mixtepec-Mixtec

- Lack of existing resources
- Lack of established linguistic description
- Related language descriptions are old, syntax based, scanned documents
- Speaker consultants work full time, often don’t have time to consistantly help edit, gloss text
- Lexical tone, adds complexity to characterization and it is not represented in the orthography
- Orthography not fully conventionalized, still changes, speakers often not aware of/don’t use the standards

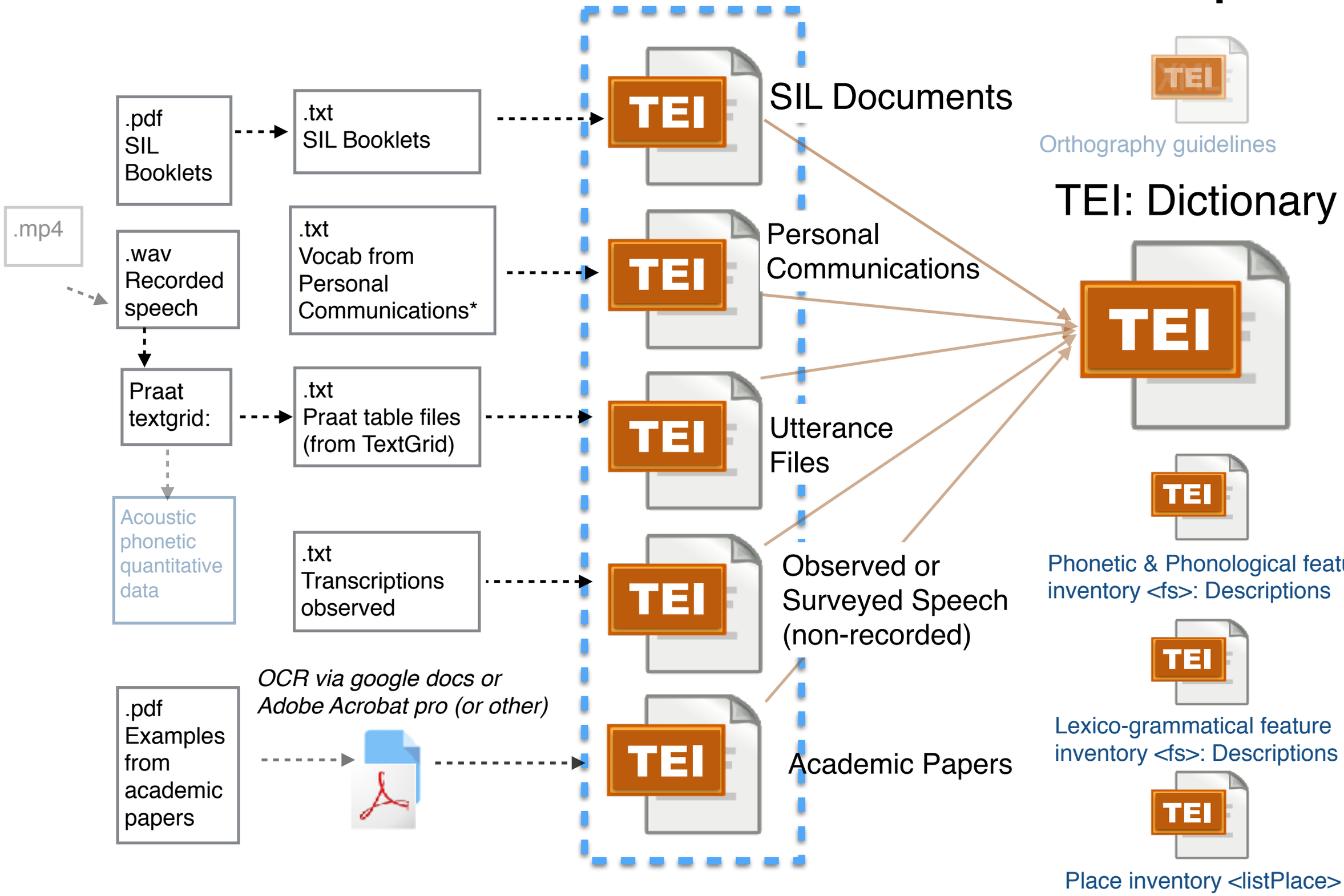
Primary Sources of Mixtepec-Mixtec Language Data

- Consultation w/ Speakers *(+- 600 recordings, written content)*
- Recordings made by speakers with other speakers
- Written content from speakers
- +-36 Children's Booklets *(Summer Institute of Linguistics Mexico)*
- Public Sources *(YouTube, etc.)*
  - Small number of papers *(phonology, some morphology)*
- Personal communications

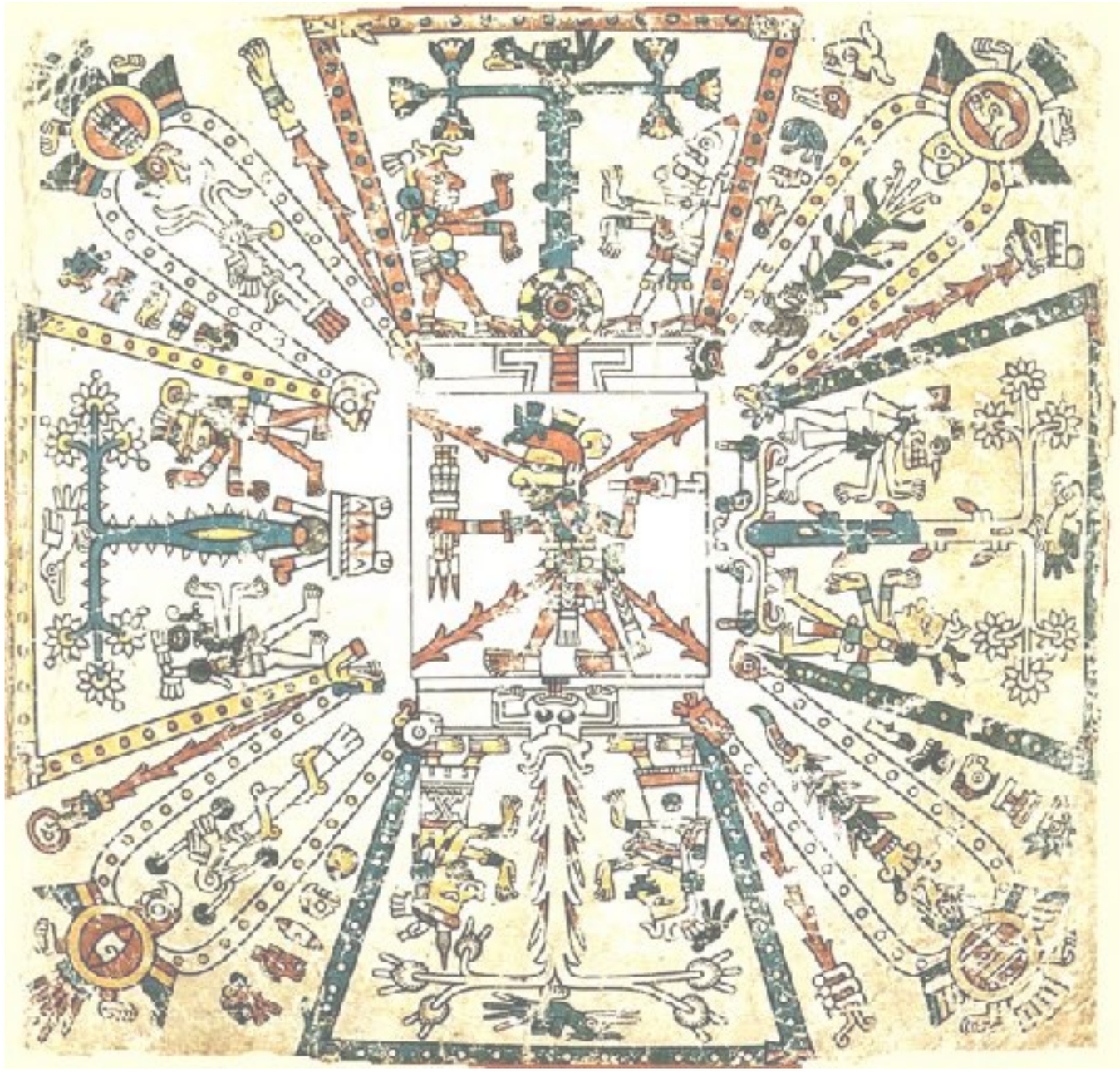
Specific TEI Output

- New Mixtec language content
- Searchable TEI corpus
- TEI dictionary
- Time aligned utterance annotated files
- Annotated TEI files of SIL booklets
- Lexical feature inventory
- Phonetic feature inventory
- Concepts inventory
- Place inventory
- Person list

Mixtec Data: Sources, Links, Output



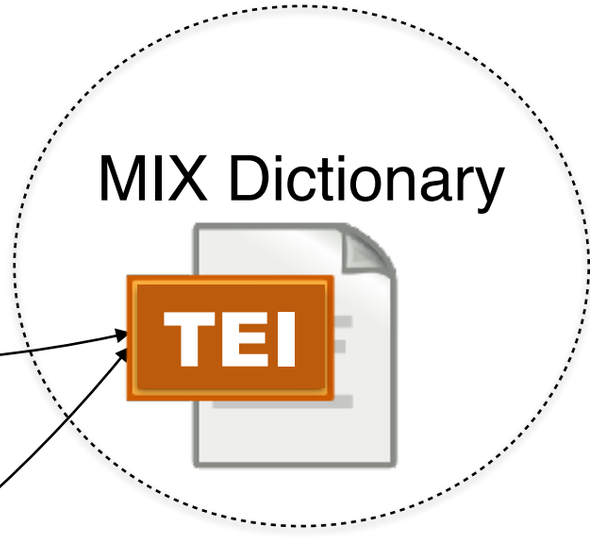
(I) Project Metadata



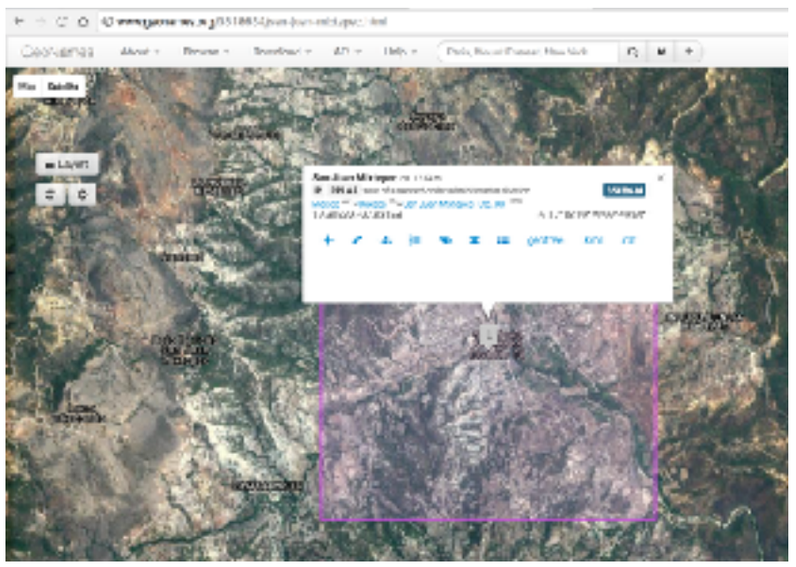
Mixtec Borgia Codex

Metadata: Places

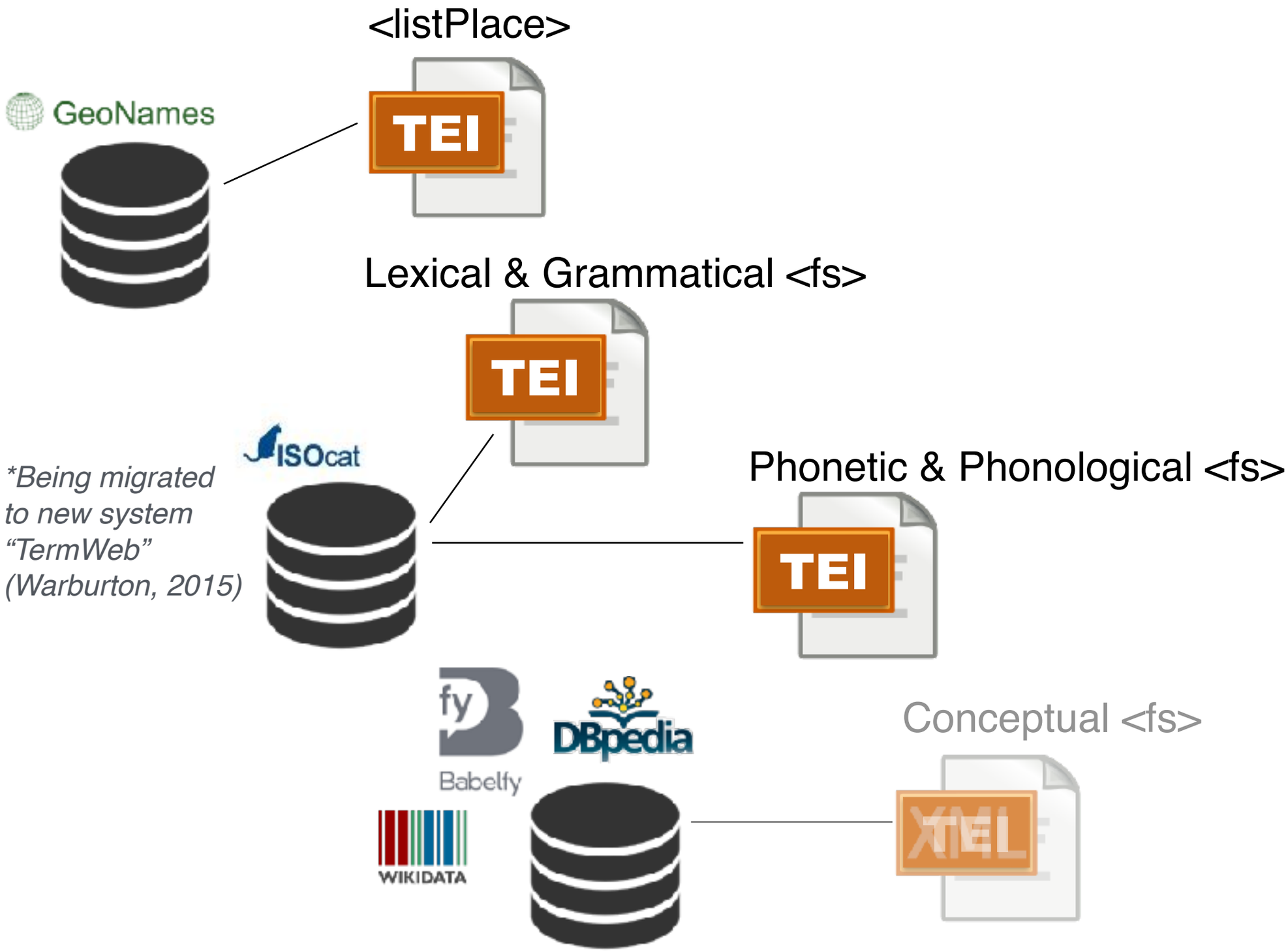
```
    <listPlace>
....
      <place xml:id="Yucunany" corresp="http://www.geonames.org/8880392">
        <placeName xml:lang="es">Yucunany</placeName>
        <placeName xml:lang="en">Yucanany</placeName>
        <placeName xml:lang="en">Yucanani</placeName>
        <placeName xml:lang="mix" cert="medium">Yukunani</placeName>
        <location>
          <geo>17.30083, -97.89389</geo>
        </location>
      </place>

      <place xml:id="SanJuanMixtepec" corresp="http://www.geonames.org/3518634">
        <placeName xml:lang="es">San Juan de Mixtepec</placeName>
        <placeName xml:lang="es">San Juan Mixtepec</placeName>
        <placeName xml:lang="mix">Snuviko</placeName>
        <placeName xml:lang="mix">Xnuviko</placeName>
        <location>
          <geo>17.30539, -97.83158</geo>
        </location>
        <note resp="JB">Mixtec place name added to geonames</note>
      </place>
....
    </listPlace>
```



Note: also included as entries in Mixtec Dictionary

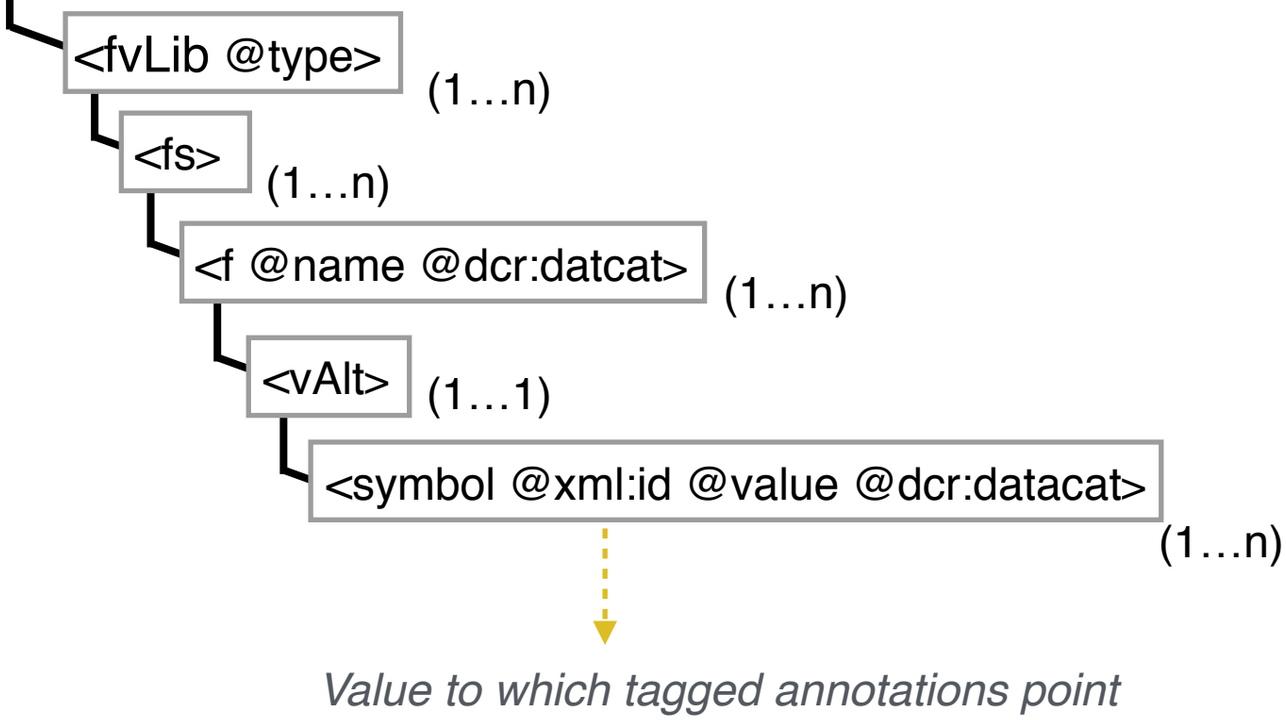


TEI Feature Structures & Standardized Resources



Linguistic Annotation: TEI Feature Structures

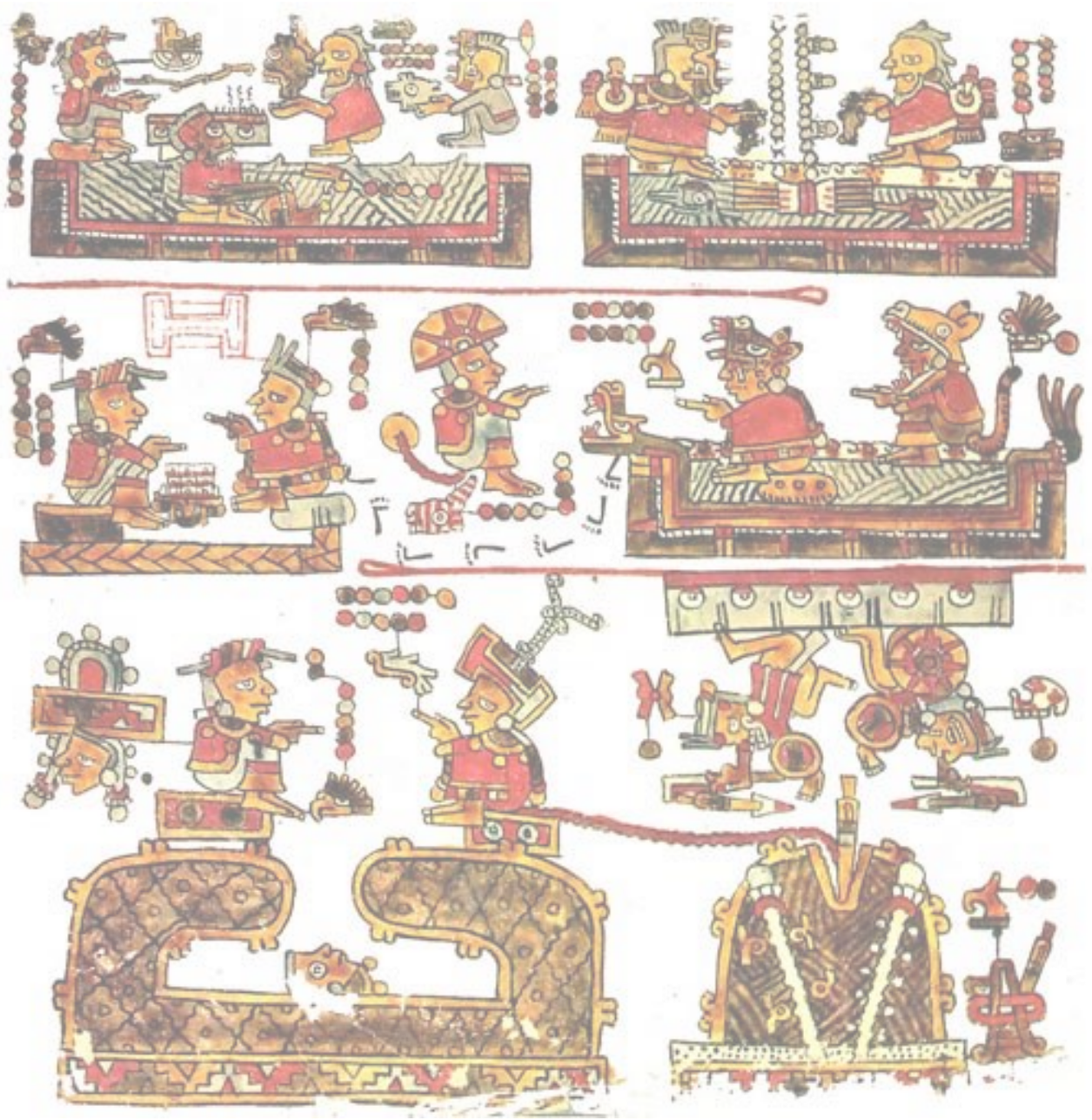
Inventory of MIX linguistic features kept in feature structures




```
<fvLib>
    <fs>
      <f name="number" xmlns:dcr="http://www.isocat.org/ns/dcr" dcr:datcat="http://www.isocat.org/datcat/DC-3351">
       <vAlt>
         <symbol xml:id="SG" value="singular" dcr:datcat="http://www.isocat.org/datcat/DC-252"/>
         <symbol xml:id="PL" value="plural" dcr:datcat="http://www.isocat.org/datcat/DC-253"/>
       </vAlt>
      </f>
    </fs>
    <!-- other feature structures here -->
  </fvLib>
```

(II) Source Documents

i. SIL Booklets



Mixtec Codex Seldon

Source Data: SIL Documents

The Summer Institute of Linguistics (SIL) documents all have an intended audience of children, there are several different document types which have different formats:

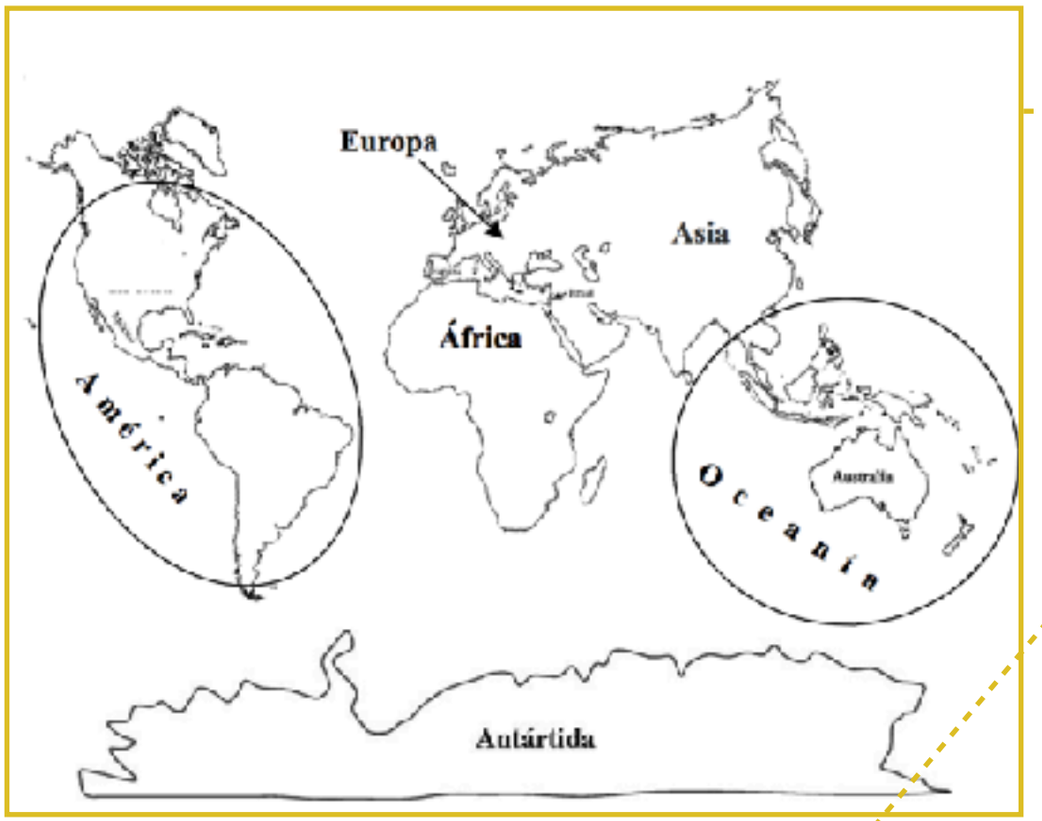
- Prose *(short stories, legends, etc.)*
- Activity/Workbooks *(picture-based excercises, crossword puzzles, mazes, etc.)*
- Vocabulary & Basic Pedigogical Reference

Current document taxonymy contains the following classifications:

- Pedagogical
  - Interactive
  - Referential
- Fiction
  - Fantasy
  - Realistic
- Folklore

SIL Documents: Prose

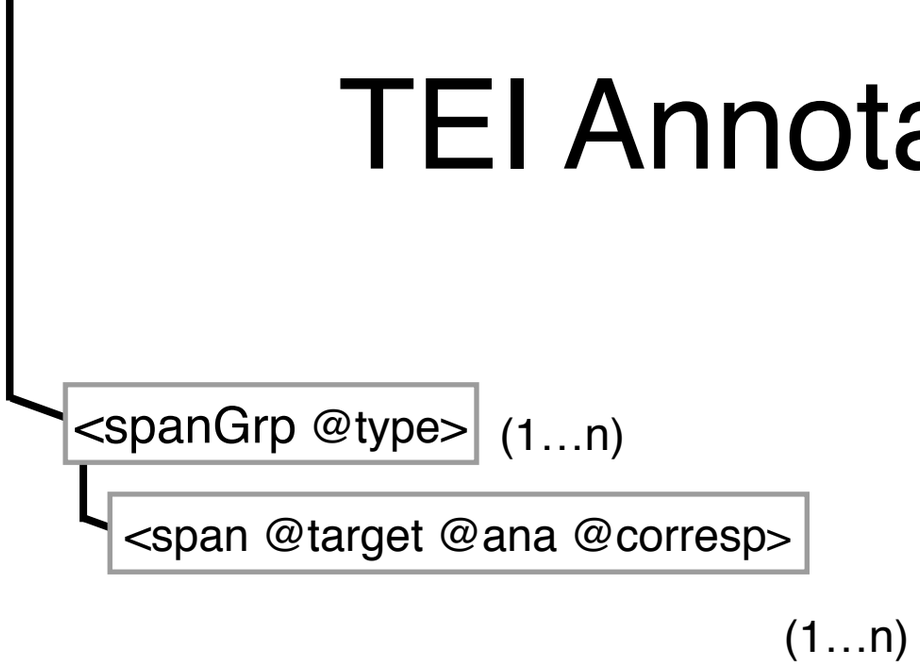
Ñu'u Ncha'i ka



Yee ñu'u tsi chikuii nuu Ñu'u Ncha'i. Yee kua'a ka chikuii cha xoo ka ñu'u. Yee ñu'u luu ka nania "islas", cha inkai ma'i chikuii. Cha ñu'u ka'nu ka nania "continente". Yee iñu "continente" nania: África, América, Antártida, Asia, Europa tsi Occanía .

```
<div xml:id="L145-13">
  <head>
    <s xml:id="L145-13-00" type="subject">
      <w xml:id="d1e1438">Ñu'u</w>
      <w xml:id="d1e1441">Ncha'i</w>
      <w xml:id="d1e1444">ka</w>
    </s>
  </head>
  <head><graphic url="L145_10.jpeg"/></head>
  <p>
    <s xml:id="L145-13-01" type="declarative">
      <w xml:id="d1e1458">Yee</w>
      <w xml:id="d1e1461">ñu'u</w>
      <w xml:id="d1e1464">tsi</w>
      <w xml:id="d1e1467">chikuii</w>
      <w xml:id="d1e1470">nuu</w>
      <w xml:id="d1e1473">Ñu'u</w>
      <w xml:id="d1e1477">Ncha'i</w>
      <pc>.</pc>
    </s>
    ....
    <s xml:id="L145-13-05" type="declarative">
      <w xml:id="d1e1555">Yee</w>
      <w xml:id="d1e1557">iñu</w><pc>"</pc>
      <w xml:id="d1e1561">continente</w><pc>"</pc>
      <w xml:id="d1e1565">nania</w> <pc>:</pc>
      <w xml:id="d1e1569">África</w><pc>,</pc>
      <w xml:id="d1e1573">América</w><pc>,</pc>
      <w xml:id="d1e1578">Antártida</w><pc>,</pc>
      <w xml:id="d1e1582">Asia</w><pc>,</pc>
      <w xml:id="d1e1586">Europa</w>
      <w xml:id="d1e1588">tsi</w>
      <w xml:id="d1e1590">Oceanía</w>
      <pc>.</pc>
    </s>
</div>
```

TEI Annotations <spanGrp>



<spanGrp> is used to annotate the following:

- Translations *(English, Spanish)*
- Grammar
- Semantics
- Etymology
- Interlinear glossed text
- General editorial notes
- (any theoretical linguistic features that fall within any of the above)

- Links annotations and translations with content
- Points to language content *(usually <w> <seg> or <s>)*
- Requires @xml:id for all values to be annotated
- Can be included within most TEI elements and thus can be inserted close to content to be annotated
- Structure and tag content correspond to project feature structure inventory <fs>

SIL Documents: Prose annotation

```
<div xml:id="L145-13">
  …
  <s xml:id="L145-13-01" type="declarative">
    <w xml:id=“d1e1458">Yee</w>
    <w xml:id="d1e1461">ñu'u</w>
    <w xml:id="d1e1464">tsi</w>
    <w xml:id="d1e1467">chikuii</w>
    <w xml:id="d1e1470">nuu</w>
    <w xml:id=“d1e1473">Ñu'u</w>
    <w xml:id=“d1e1477">Ncha'i</w>
    <pc>.</pc>
  </s>
  …
</div>
```

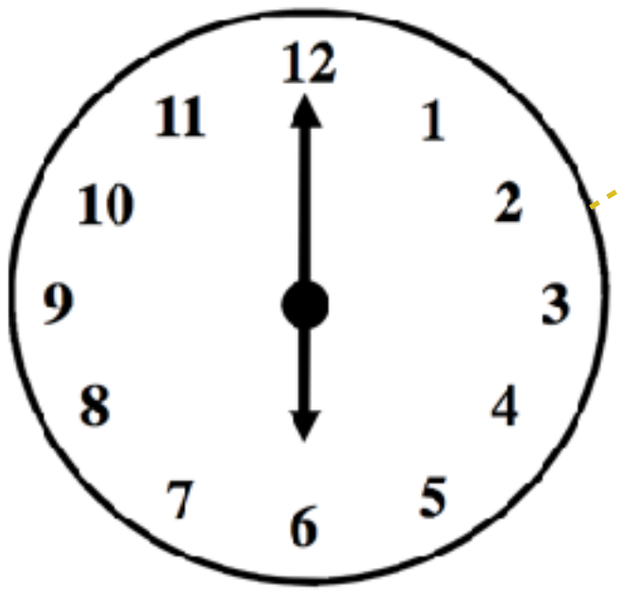
Annotations: Translations

```
<spanGrp type=“translation">
  <span target="#d1e1458" xml:lang="en">there is</span>
  <span target="#d1e1458" xml:lang="es">hay</span>
  <span target="#d1e1461" xml:lang="en">land</span>
  <span target="#d1e1461" xml:lang="es">tierra</span>
  <span target="#d1e1464" xml:lang="en">and</span>
  <span target="#d1e1464" xml:lang="es">y</span>
  <span target="#d1e1467" xml:lang="en">water</span>
  <span target="#d1e1467" xml:lang="es">agua</span>
  <span target="#d1e1470 #d1e1473 #d1e1477" xml:lang="en">on Earth</span>
  <span target="#d1e1470 #d1e1473 #d1e1477" xml:lang="es">en la tierra</span>
  <span target="#d1e1473 #d1e1477" xml:lang="en">Earth</span>
  <span target="#d1e1473 #d1e1477" xml:lang="es">la tierra</span>

  <span target="#L145-13-01"xml:lang="en">There is land and water on the Earth.</span>
  <span target="#L145-13-01" xml:lang="es">Hay tierra y agua en la Tierra.</span>

</spanGrp>
```

Annotations: Sense (Concepts)

```
<spanGrp type="sense">
  <span target=“#d1e1473 #d1e1477” corresp="http://dbpedia.org/resource/Earth"/>
  <span target="#d1e1467" corresp="http://dbpedia.org/resource/Water"/>
</spanGrp>
```

SIL Documents: Workbook
(reference version w/answers)



¿Nchii hora kui?

Ka iñu ntaa.

3

```
<div xml:id="L093-01">
  <head>
    <graphic url="L093-1-what_time_is_it-6.jpg"/>
  </head>
  <label>
    <time>6:00</time>
  </label>
  <lb/>
  <p>
    <s xml:id="d1e160" type=“interrogative">
      <pc>¿</pc>
      <w xml:id="d1e163">Nchii</w>
      <w xml:id="d1e165">hora</w>
      <w xml:id="d1e167">kui</w>
      <pc>?</pc>
    </s>
    <lb/>
    <s xml:id="d1e174" type=“declarative">
      <w xml:id=“d1e175">Kaa</w>
      <w xml:id="d1e177">iñu</w>
      <w xml:id="d1e179">ntaa</w>
      <pc>.</pc>
    </s>
  </p>
</div>
```

SIL Documents: Workbook

(*reference version w/answers*) annotation

```
<div xml:id="L093-01">
.....
 <p>
    <s xml:id="d1e160" type=“interrogative">
      <pc>¿</pc>
      <w xml:id="d1e163">Nchii</w>
      <w xml:id="d1e165">hora</w>
      <w xml:id=“d1e167">kui</w>
      <pc>?</pc>
    </s>
    <lb/>
    <s xml:id="d1e174" type=“declarative">
      <w xml:id=“d1e175">Kaa</w>
      <w xml:id="d1e177">iñu</w>
      <w xml:id="d1e179">ntaa</w>
      <pc>.</pc>
    </s>
  </p>
</div>
```

Annotations: Interlinear Glossed Text

```
<spanGrp type=“igt” target=“#d1e160">
  <span target=“#d1e163">wh</span>
  <span target="#d1e165">time</span>
  <span target="#d1e167">cop-incmpl;3s</span>
</spanGrp>
```

```
<spanGrp type=“igt” target=“#d1e174">
  <span target=“#d1e175”>cop-eqtv</span>
  <span target="#d1e177">six</span>
  <span target=“#d1e179”>o'clock</span>
</spanGrp>
```

Annotations: Grammar

```
<spanGrp type="gram">
  <span type="sentence" target="#d1e160" ana="#Q #WH #TEMP"/>
  <span type="phrase" target="#d1e163 #d1e165" ana="#ADVP #WH #TEMP"/>
  <span type="pos" target="#d1e167" ana="#COP #INCMPL”/>
  <span type="aspect" target="#d1e167" ana=“#INCMPL"/>
  <span type="person" target="#d1e169" ana=“#3PERS"/>
  <span type="number" target="#d1e169" ana=“#SG"/>
</spanGrp>
```

SIL Documents: Basic Vocabulary

chumi xini ka'nu
tecolote
búho cornado

chumi lunchi
tecolote llanero
tecolote zancón

chumi sai
tecolotito

```
<item>
    <graphic url="Aves-01.png"/>
    <seg xml:id="d1e35" xml:lang="mix" type="compound">
      <w xml:id="d1e36">chumi</w> <w xml:id="d1e38">lunchi</w>
    </seg>
    <seg xml:id="d1e40" xml:lang="es" type="compound">
      <w xml:id="d1e41">tecolote</w> <w xml:id="d1e43">llanero</w>
    </seg>
    <seg xml:id="d1e45" xml:lang="es" type="compound">
      <w xml:id="d1e46">tecolote</w> <w xml:id="d1e48">zancón</w>
    </seg>
  </item>
  <item>
    <graphic url="Aves-02.png"/>
    <seg xml:id="d1e53" xml:lang="mix">
      <w xml:id="d1e54">chumi</w> <w xml:id="d1e56">xini</w>  <w xml:id="d1e58">ka'nu</w>
    </seg>
    <seg xml:id="d1e60" xml:lang="es">
      <w xml:id="d1e61">tecolote</w>
    </seg>
    <seg xml:id="d1e63" xml:lang="es" type="compound">
      <w xml:id="d1e64">búho</w> <w xml:id="d1e66">cornado</w>
    </seg>
</item>
<item>
    <graphic url="Aves-03.png"/>
    <seg xml:id="d1e71" xml:lang="mix" type="compound">
      <w xml:id="d1e72">chumi</w> <w xml:id="d1e74">sai</w>
    </seg>
    <seg xml:id="d1e76" xml:lang="es">
      <w xml:id="d1e77">tecolotito</w>
    </seg>
</item>
```

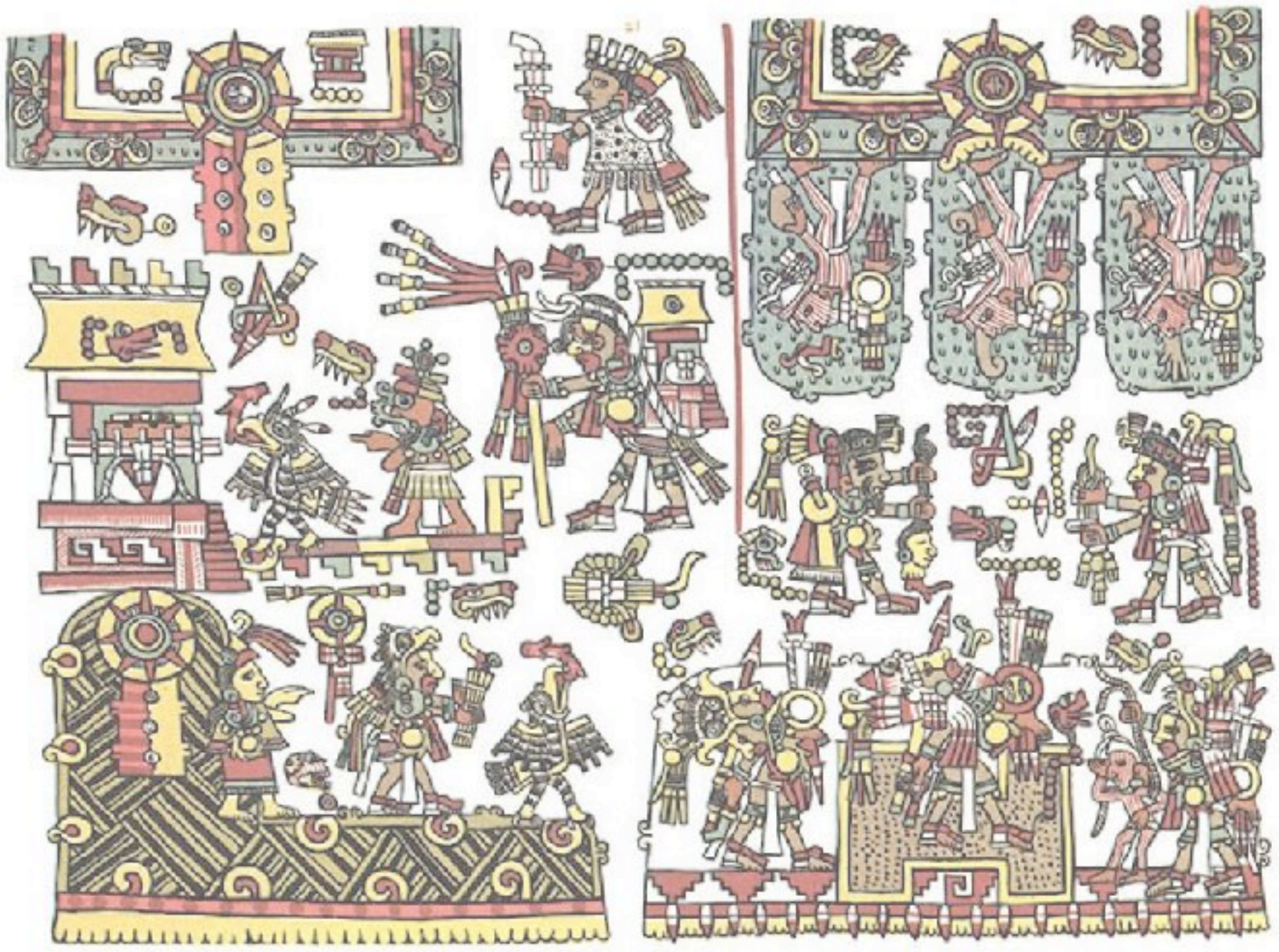
SIL Documents: Basic Vocabulary Annotation

```
<item>
    <graphic url="Aves-02.png"/>
    <seg xml:id="d1e53" xml:lang="mix" type="compound">
      <w xml:id="d1e54">chumi</w>
      <w xml:id="d1e56">xini</w>
      <w xml:id="d1e58">ka'nu</w>
    </seg>
    <seg xml:id="d1e60" xml:lang="es-MEX">
      <w xml:id="d1e61">tecolote</w>
    </seg>
    <seg xml:id="d1e63" xml:lang="es" type="compound">
      <w xml:id="d1e64">búho</w>
      <w xml:id="d1e66">cornado</w>
    </seg>
</item>
```

Annotations: Translations; Sense (concept); Lexical Relations

```
<linkGrp type="translation">
    <link target="#d1e53 #d1e60"/>
    <link target="#d1e53 #d1e63"/>
</linkGrp>
<spanGrp type="sense"><!-- concept -->
    <span target="#d1e53 #d1e60 #d1e63" corresp="http://dbpedia.org/page/Great_horned_owl"/>
</spanGrp>
<spanGrp type="lexicalRelations"><!-- Regional differences marked though lang tag (BCP47 -->
    <span type="synonym" target="#d1e46 #d1e48"/>
</spanGrp>
```

(II) Source Documents
ii. Spoken Language Resources

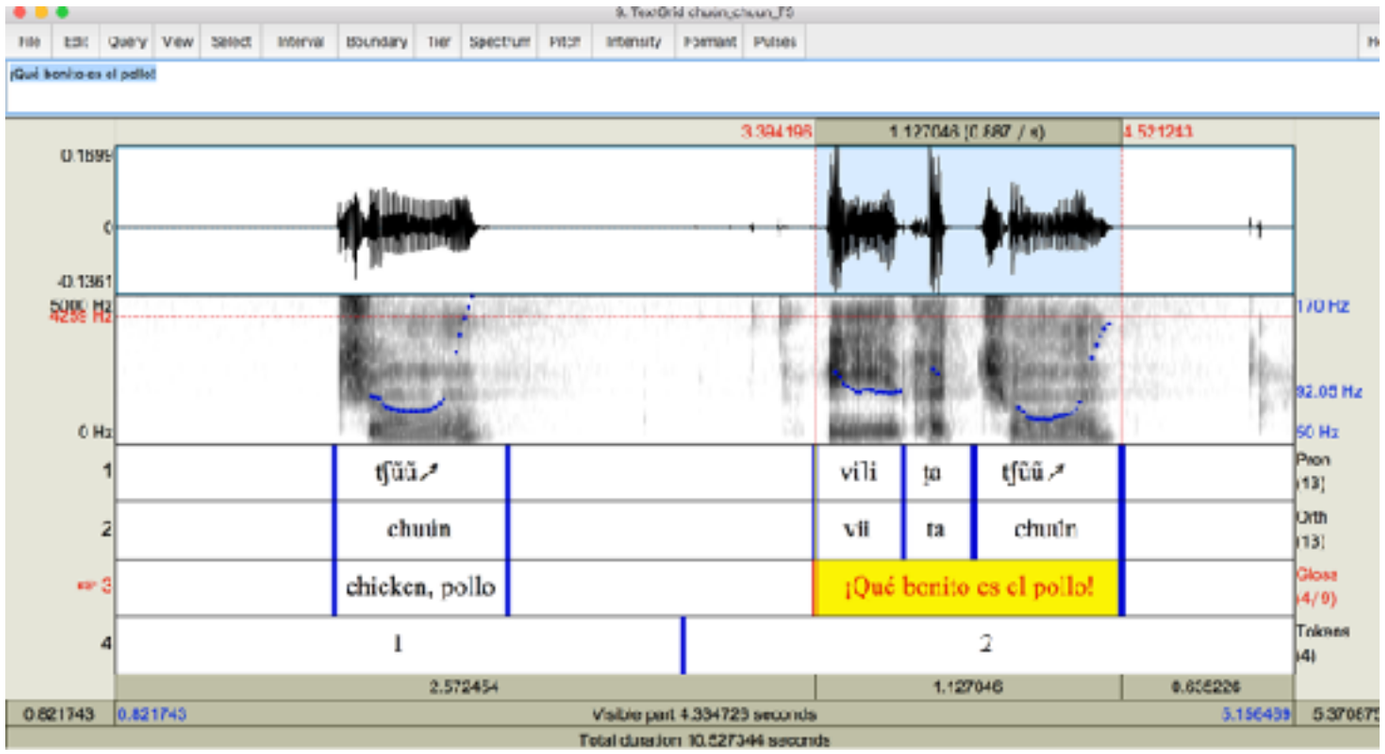


Codex Zouche-Nuttall, British Museum.

Speech Annotation: Toolkits & Features

	Praat	Exmaralda
metadata*	no	yes
spectrogram view	yes	no
XML/TEI output option	no	yes
tiered/ time aligned segmentation	yes	yes
scripting	yes	no
TEI/XML export	no	yes
corpus managment, searching	*(via scripting)*	yes *(text based only)*
video annotation	no	yes
visualization	yes*	yes
quantitivative data extraction	yes	no
pitch (F0) view/analysis	yes	no

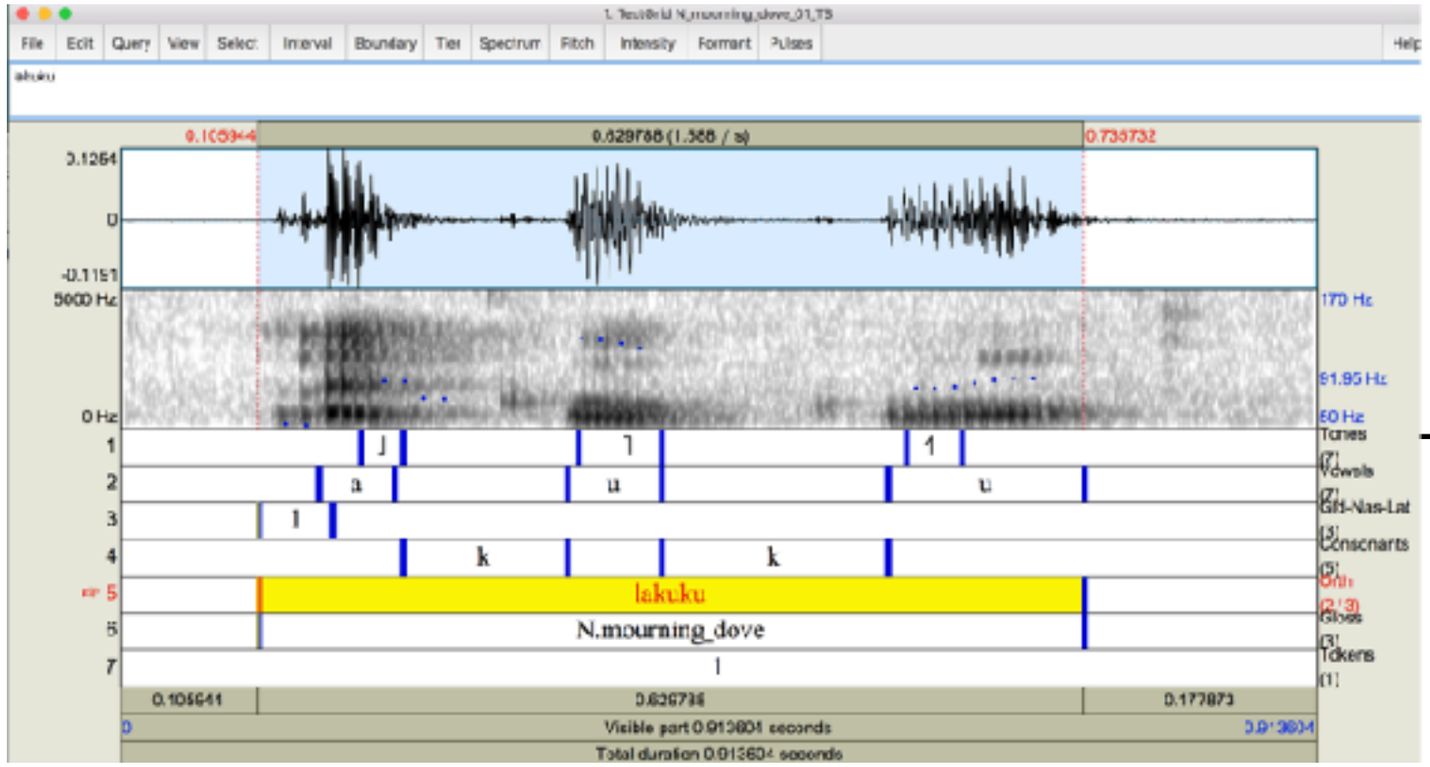
Speech Annotation: Praat
(basic transcription method)



tmin	tier	text	tmax	
0	Tokens	1	2.91	
1.63	Gloss	chicken, pollo		2.26
1.63	Pron	tʃũũ↗	2.26	
1.63	Orth	chuún	2.26	
2.91	Tokens	2	5.18	
3.39	Orth	vii	3.72	
3.39	Pron	viĩi	3.72	
3.39	Gloss	¡Qué bonito es el pollo!		4.52
3.72	Pron	t̪a	3.98	
3.72	Orth	ta	3.98	
3.98	Orth	chuún	4.52	
3.98	Pron	tʃũũ↗	4.52	

Utterance File <u>

Speech Annotation: Praat
(phonetic focus transcription)

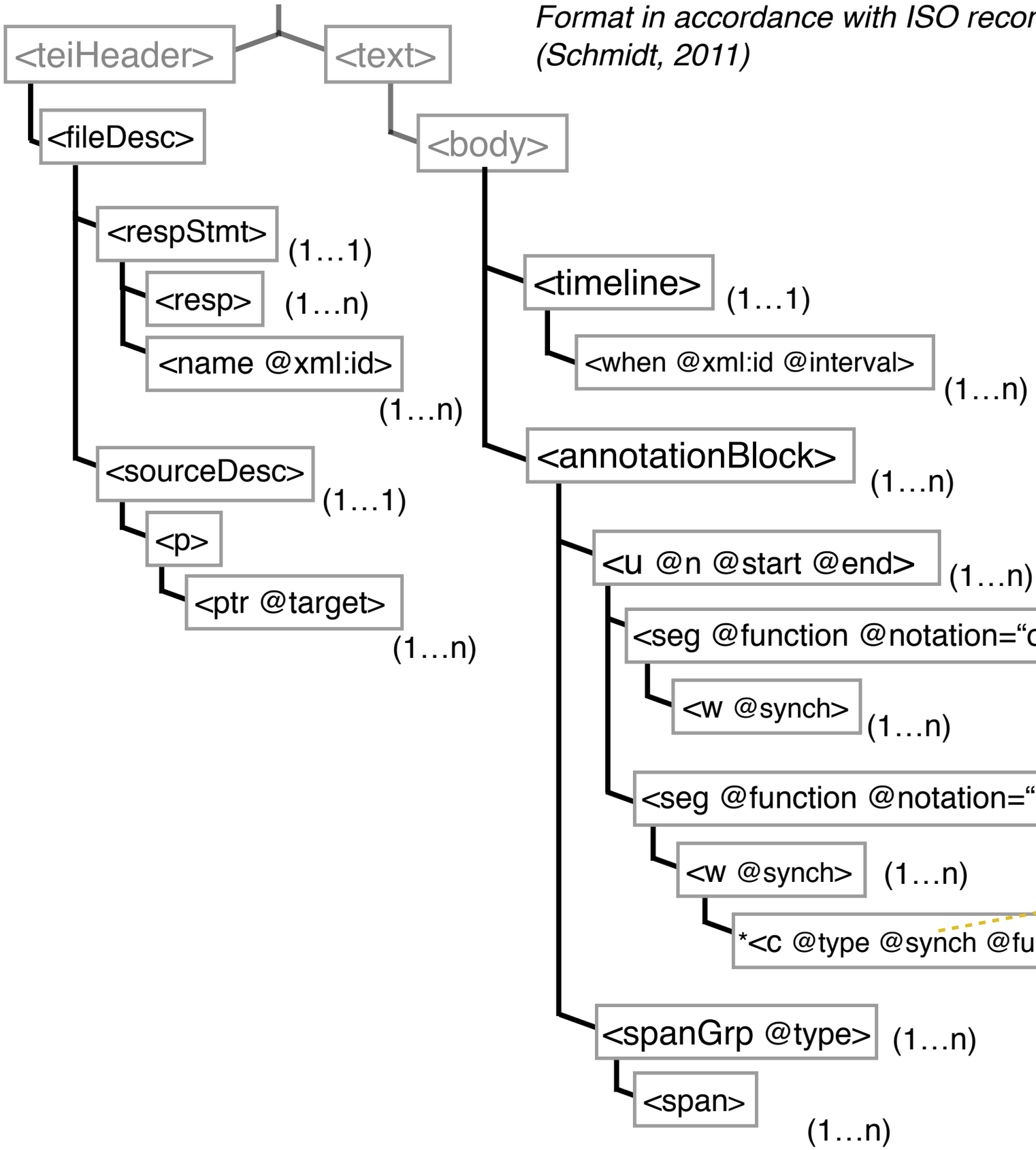


tmin	tier	text	tmax	
0	Tokens	1	0.91	
0.11	Gld-Nas-Lat	l	0.16	
0.11	Orth	lakuku	0.74	
0.11	Gloss	N.mourning_dove		0.74
0.15	Vowels	a	0.21	
0.18	Tones	˩	0.22	
0.22	Consonants	k	0.34	
0.34	Vowels	u	0.41	
0.35	Tones	˥	0.41	
0.41	Consonants	k	0.59	
0.59	Vowels	u	0.74	
0.60	Tones	˧	0.64	

Utterance File <u>

TEI Utterance files (from Praat)

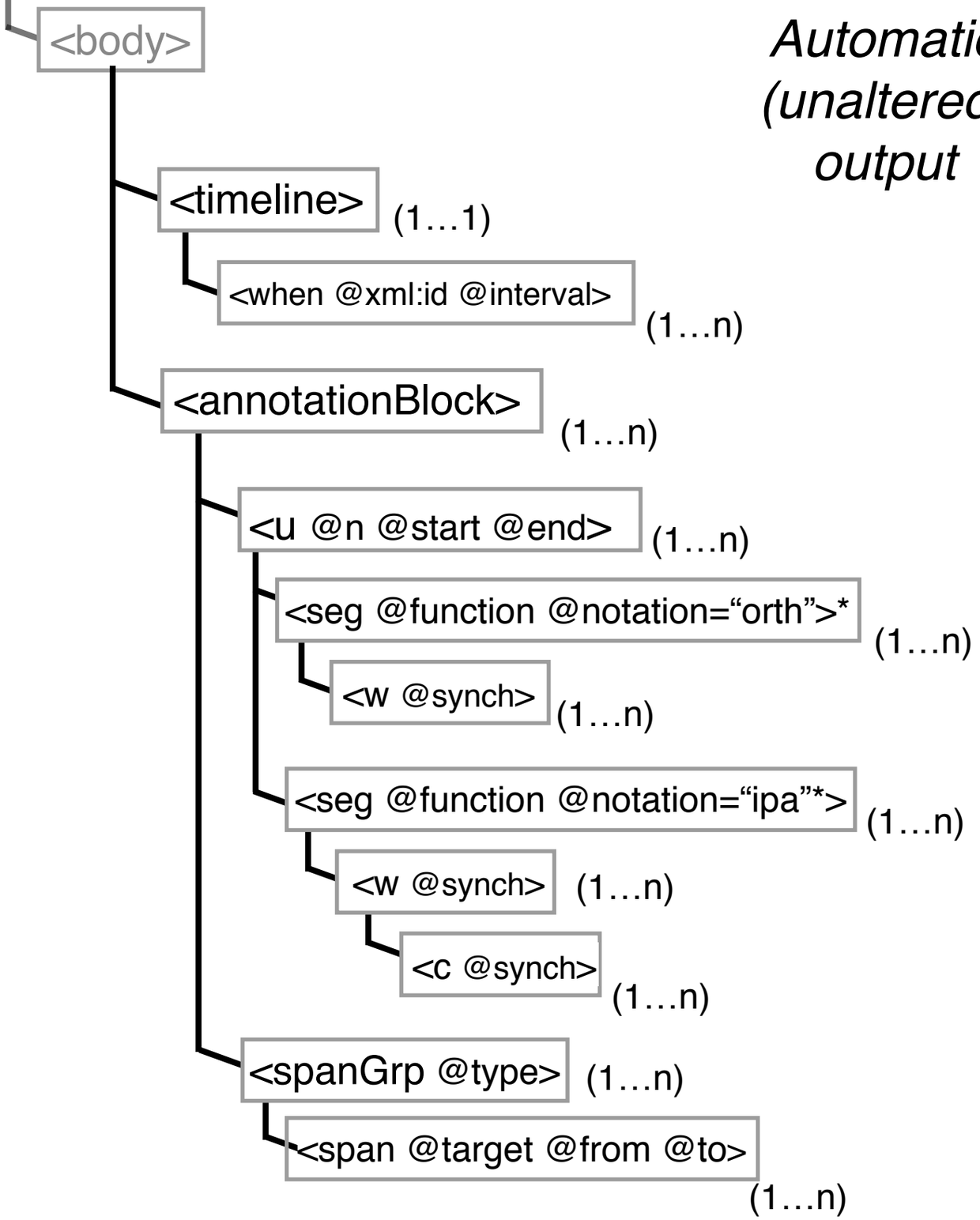
Format in accordance with ISO recommendation for speech transcription (Schmidt, 2011)



- One utterance file per Praat TextGrid
- Source .wav and praat textgrid filenames in header <ptr @target> within <sourceDesc>
- Can generate speaker info in header from file name <respStmt>

<c>'s correspond to <fs> values for phonetic/phonological invetory (only included in output from fully segmented (phonetic focus) praat annotations)

TEI Utterance files (from Praat)



Automatic (unaltered) output

```
<body>
  <timeline>
    <when xml:id="T1" interval="0.11"/>
    <when xml:id="T2" interval="0.15"/>
    <when xml:id="T3" interval="0.18"/>
    <when xml:id="T4" interval="0.22"/>
    <when xml:id="T5" interval="0.34"/>
    <when xml:id="T6" interval="0.35"/>
    <when xml:id="T7" interval="0.41"/>
    <when xml:id="T8" interval="0.59"/>
    <when xml:id="T9" interval="0.60"/>
    <when xml:id="T10" interval="0.74"/>
  </timeline>
  <annotationBlock>
    <u xml:id="d1e39" n="1" start="0" end="0.91">
      <seg xml:id="d1e40" function="utterance" notation="orth">
        <w xml:id="d1e41" synch="#T1">lakuku</w>
      </seg>
      <seg xml:id="d1e44" function="utterance" notation="ipa">
        <w xml:id="d1e45" synch="#T1">
          <c>l</c>
          <c>a</c>
          <c function="tone">˩</c>
          <c>k</c>
          <c>u</c>
          <c function="tone">˥</c>
          <c>k</c>
          <c>u</c>
          <c function="tone">˩</c>
        </w>
      </seg>
    </u>
    <spanGrp type="praatGloss">
      <span from="#T1" to="#T10">N.mourning_dove</span>
    </spanGrp>
    .....
  </annotationBlock>
</body>
```

TEI Utterance files (from Praat):Annotated

```
<timeline>
……
</timeline>
<annotationBlock>
  <u xml:id="d1e39" n="1" start="0" end="0.91">
    <seg xml:id="d1e40" function="utterance" notation="orth">
      <w xml:id="d1e41" synch="#T1">lakuku</w>
    </seg>
    <seg xml:id="d1e44" function="utterance" notation="ipa">
      <w xml:id="d1e45" synch="#T1">
        <c>l</c>
        <c>a</c>
        <c function="tone">˩</c>
        <c>k</c>
        <c>u</c>
        <c function="tone">˥</c>
        <c>k</c>
        <c>u</c>
        <c function="tone">˥</c>
      </w>
    </seg>
  </u>
  <spanGrp type="praatGloss">
    <span from="#T1" to="#T10">N.mourning_dove</span>
  </spanGrp>
  <spanGrp type="gram">
    <span type="pos" target="#d1e41 #d1e45" ana="#N"/>
  </spanGrp>
  <spanGrp type="semantics">
    <span type="sense" target="#d1e41 #d1e45" corresp="http://dbpedia.org/resource/Mourning_dove"/>
    <!-- is_a:Bird -->
    <span type="domain" target="#d1e41 #d1e45" corresp="http://dbpedia.org/resource/Bird"/>
  </spanGrp>
  <spanGrp type="translation">
    <span target="#d1e41 #d1e45" xml:lang="en" corresp="https://en.wiktionary.org/wiki/mourning_dove">mourning dove</spa
    <span target="#d1e41 #d1e45" xml:lang="es" corresp="https://es.wiktionary.org/wiki/tortolita">tortolita</span>
  </spanGrp>
</annotationBlock>
```

to TEI Dictionary: (value of) //form[@type="lemma"]/orth

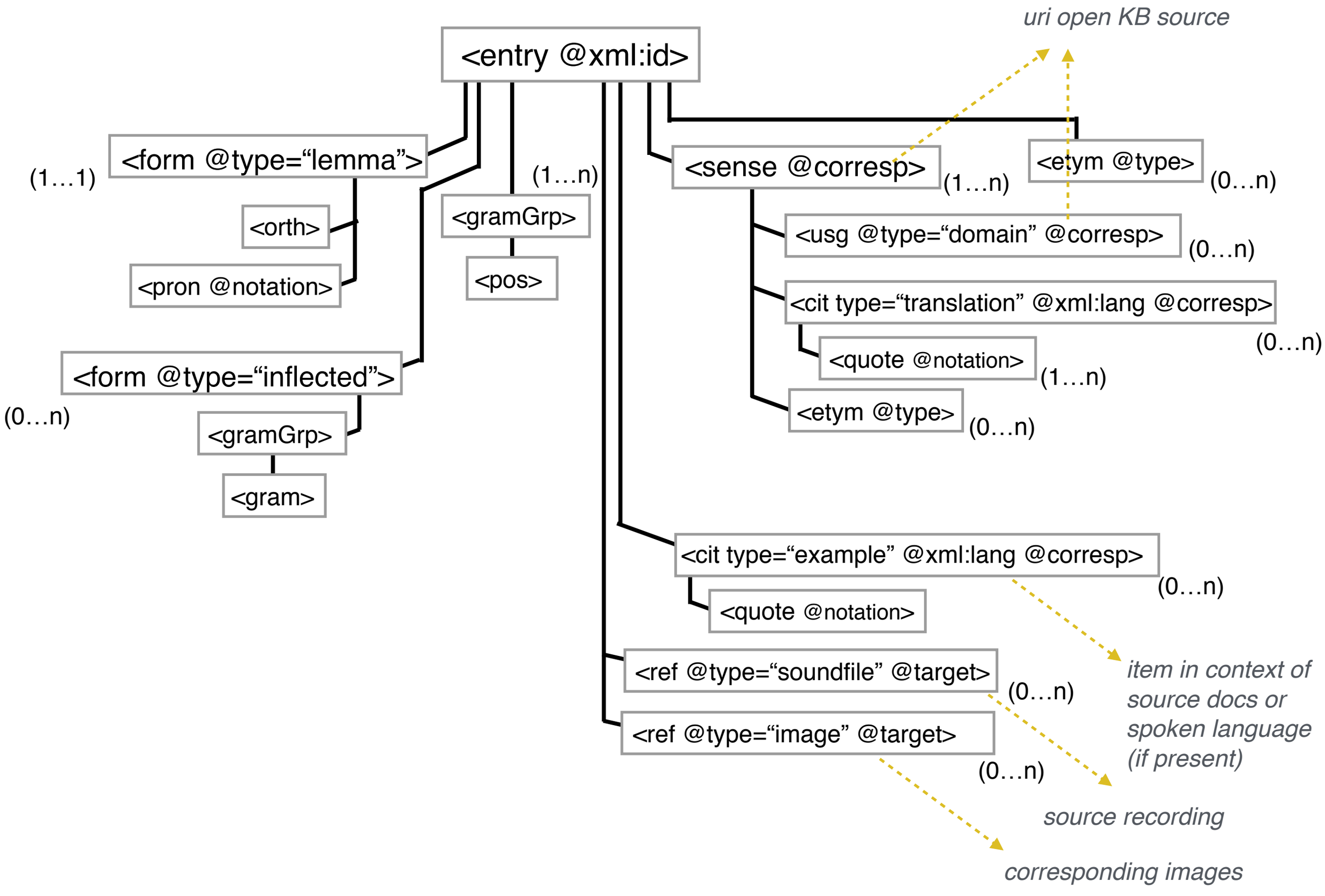
to Dictionary: (value of) //form[@type="lemma"]/pron[notation="ipa"]

Manually added in Oxygen

(III) TEI Dictionary

Mixtec Codex Nuttal- British Museum

TEI Dictionary Structure



TEI Dictionary Entry: Basic example

```
<entry xml:id="bird-mourning_dove">
   <form type="lemma">
     <orth>lakuku</orth>
     <pron notation="“ipa”">laɻkuʔku˥</pron>
      <!-- include each unique pronunciation (at least until incidental outliers are identifiable)-->
   </form>
     <gramGrp>
       <pos>noun</pos>
     </gramGrp>
   <sense corresp="“http://dbpedia.org/resource/Mourning_dove">
     <usg type="domain" corresp="http://dbpedia.org/resource/Bird" xml:lang="mix">Saa</usg>
    <cit type="translation" xml:lang="en" corresp="“https://en.wiktionary.org/wiki/mourning_dove">
      <oRef>mourning dove</oRef>
    </cit>
    <cit type="translation" xml:lang="es" corresp="“https://es.wiktionary.org/wiki/tortolita">
      <oRef>tortolita</oRef>
    </cit>
  </sense>
  <cit type="example" corresp="“/SIL_docs/L152/L152-tok.xml#L152-01-01”>
    <quote>Iin kii ra iin <oRef>lakuku</oRef> kunia tanta'i tsi iin ncho'o, cha koo xu'in sa'i viko.</quote>
  </cit>
  <ref type="soundfile" target="N_mourning_dove_01_TS.wav"/>
  <!-- could also include references to images (where available) -->
</entry>
```

(III) TEI Dictionary
ii. Etymology

Mixtec Codex Bodley

TEI Dictionary Etymology

Bowers & Romary (2016) propose expansion and refinement of etymology section of the TEI dictionary module to include detailed proposals for the encoding of many important processes of linguistic change; e.g.

- Sense changes:
  - Metaphor
  - Metonymy
  - Grammaticalization *(and sub-process)*
  - *(others)*
- Compounding
- Phonetic changes (any)
- Borrowing
- Inheritance

TEI Dictionary Entry: Etymological Markup

TEI etymology markup format as per Bowers & Romary (2016)

```
<entry xml:id="kidney" xml:lang="mix">
  <form type="lemma">
    <orth>ntuchi</orth>
    <pron notation="ipa">nduɭʧi/</pron>
     <gramGrp>
      <pos>noun</pos>
     </gramGrp>
  </form>
  <sense corresp="http://dbpedia.org/resource/Kidney">
    <usg type="dom"corresp="http://dbpedia.org/resource/Human_body">Body</usg>
    <usg type="dom" corresp="http://dbpedia.org/resource/Human_organs">InternalOrgans</usg>

     <etym type="metaphor">
      <cit type="etymon">
        <oRef corresp="#bean">ntuchi</oRef>
        <pRef notation="ipa" corresp="#bean">nduɭʧi/</pRef>
        <ref type="sense" corresp="http://dbpedia.org/resource/Bean"/>
        <usg type="dom" corresp="http://dbpedia.org/resource/Category:Edible_legumes">Legume</usg>
        <gloss>bean</gloss>
      </cit>
     </etym>

     <cit type="translation" xml:lang="en">
      <oRef>kidney</oRef>
     </cit>
  </sense>
</entry>
```

Next Steps

- Make use of/ implement the @lemma in <w> to link all inflected word forms/phrases with their common lemma
- Implement First Order Logic-Based linguistic structual descriptions
- Establish more refined translation typology
- Improve/standardize automatic processing, markup programming
- Disseminate the corpus in CC-BY
- Produce corpus based studies of polysemy and etymological processes (particularly in Body-part terms)

Inria
INVENTEURS DU MONDE NUMÉRIQUE

ÖAW
AUSTRIAN ACADEMY OF SCIENCES