Modules/Tables

Etyma
Reconstructions (PTB + meso-languages). Each reconstruction has a unique id (‘tag number’). Meso-reconstructions in this table have no higher-level PTB reconstructions (yet).

Lexicon
Transcriptions of lexical items imported from various sources (dictionaries, grammars, other scholarly works, contributions from fieldworkers, etc.). A lexical item that supports a particular reconstruction is linked to the etymon by its tag number.

Language Names (languagenames)
Metadata for all languages in the database.

Source Bibliography (srcbib)
Metadata for all sources of data.

Chapters
Semantic hierarchy in tabular format.

Glosswords
Glosses associated with each semantic category (‘semkey’) (used for automatic semantic categorization).

HPTB
Roots from the Handbook of Proto-Tibeto-Burman (Matisoff 2003). Useful as an electronic index of the reconstructions in that work.

Morphemes
(under development) Prototype table containing algorithmically-extracted morphemes from lexical items.
Sound Laws (soundlaws)
(under development) Prototype table containing sound correspondences and regular sound changes extracted from tagged cognate sets.

Language Groups
The Tibeto-Burman subgrouping scheme used in the database.

Mesoroots
All meso-level reconstructions associated with higher-level PTB reconstructions.

Etymologies
All STEDT-approved associations between etyma and supporting forms. (Allows you to search on the properties of both at the same time, e.g. find all Tibetan lexical items whose PTB etyma contain *-añ.)

Exploring the Semantic Hierarchy

  Chapter browser
  Hierarchy of semantic categories and etyma under each category.

  Semkeys and glosswords
  Hierarchy of semantic categories and associated glosses (for automatic categorization).

  Grid view
  Semantic hierarchy in volume-fascicle grid view.
Searching

Search Fields
(fields most commonly searched are listed first)

Lexicon

gloss: keyword search of gloss field
form: regular expression (POSIX ERE) search of transcribed lexical item
language: prefix search of language name
grp (language group): searches a particular subgroup
   - strict option restricts the search to top-level node of subgroup

Other lexicon search fields
rn (record number): unique numeric id of lexical item
analysis: all stedt & user tags
[username]ʼs analysis: all tags by currently-selected user
gfn: regular expression search of grammatical function (where provided by the source)
srcabbr (source abbreviation): prefix search of source abbreviation
srcid (source id): page number/set number/etc. for locating item in original source
semkey (semantic key): position in semantic hierarchy (e.g. “1.2.7”)
lgid (language id): unique numeric id for source+language combination
lgcode (language code): numeric code for language (under development)

Etyma

protogloss: keyword search of proto-gloss field
protoform: regular expression search of reconstructed form
plg (proto-language): restricts search to particular proto-language

Other etyma search fields
# (tag number): unique numeric id of etymon
reflexes: searches etyma by number of supporting forms
ch. (chapter): position in semantic hierarchy/thesaurus (e.g. “1.2.7”)
seq (sequence number): allofam group / print sequence number
tagging note: regular expression search of notes field
status: regular expression search of status field
notes: searches etyma by number of etymon notes
public: restricts search based on STEDTʼs confidence in etymon
owner: restricts search to etyma owned by a particular user
Languagenames

- **language**: prefix search of language name
- **silcode**: searches ISO 639-3 codes
- **grp** (language group): searches a particular subgroup
  - **strict** option restricts the search to top-level node of subgroup

Other languagenames search fields

- **lgid** (language id): unique numeric id for source+language combination
- **srcabbr** (source abbreviation): prefix search of source abbreviation
- **lgcode** (language code): numeric code for language
  - under development
- **notes**: regular expression search of notes field

Srbib

- **author**: regular expression search of authors
- **year**: regular expression search of year of publication/submission
- **title**: regular expression search of titles

Other srcbib search fields

- **srcabbr** (source abbreviation): prefix search of source abbreviation

Special Search Characters

**Word boundary:** \b

*Function*: Adds a word boundary to the search term

*Applies to*: Any regular expression search field, e.g. form (lexicon), protoform (etyma)

*Example*: `ta\b` searches for `ta` followed by word boundary (space, dash, end of field, etc.)

<table>
<thead>
<tr>
<th>Results</th>
<th>Search</th>
<th>ta</th>
<th>tab</th>
</tr>
</thead>
<tbody>
<tr>
<td>ata</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ta-sak</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>tak</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

**Suppress default word boundary:** * (at beginning of field)

*Function*: Cancels the surrounding word boundaries inserted by default in (proto)gloss fields

*Applies to*: gloss (lexicon), protogloss (etyma), srcabbr (lexicon, languagenames, srcbib)

*Example*: `*thin` searches for all instances (even word-internal) of the character sequence `thin`
Tagged/untagged records: 0 (untagged), !0 (tagged)

*Function*: Restricts search to (un)tagged records

*Applies to*: analysis fields (lexicon)

*Example*: The results of a lexicon search for the gloss ‘dog’ change depending on the search term in the analysis field:

<table>
<thead>
<tr>
<th>Analysis search term</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(blank)</td>
<td>all lexicon records with gloss containing ‘dog’</td>
</tr>
<tr>
<td>0</td>
<td>all ‘dog’ lexicon records with no etymon tag (from any user)</td>
</tr>
<tr>
<td>!0</td>
<td>all ‘dog’ lexicon records with an etymon tag (from any user)</td>
</tr>
</tbody>
</table>

Boolean operators: & (AND), , (OR), ! (NOT)

*Function*: Searches for records matching both terms (AND), either term (OR), or excluding a particular term (NOT)

*Applies to*: Nearly any search field

*Example*:

<table>
<thead>
<tr>
<th>Lexicon gloss search terms</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>all lexicon records with gloss containing ‘dog’</td>
</tr>
<tr>
<td>bark, dog</td>
<td>all lexicon records with gloss containing ‘dog’ or ‘bark’</td>
</tr>
<tr>
<td>bark&amp;dog</td>
<td>all lexicon records with gloss containing both ‘dog’ and ‘bark’</td>
</tr>
<tr>
<td>bark&amp;!dog</td>
<td>all lexicon records with gloss containing ‘bark’ but not ‘dog’</td>
</tr>
</tbody>
</table>
Content Creation
(Requires account with tagger-level privileges)

Contributing Data (lexical items / reconstructions)
Email your data file to stedt@berkeley.edu and we’ll import it for you.

Annotating
Various objects in the database can have notes attached to them:

Lexical items
Accessible via lexicon table or etymon view

Adding a lexicon note:
Click [+] in notes field of lexicon record.
Type note in text box, using markup guidelines for special formatting, (Note type will be ‘Internal’.)
Click ‘Add Note’.
A footnote reference will appear in the notes field, linking to the note text at the bottom of the page.

Editing a lexicon note:
Click footnote link in notes field to jump down to note.
Click ‘Edit’ button next to note.
Make desired edits in the text box on the far right. Markup guidelines can be toggled on & off by clicking ‘markup’.
Click ‘Save Note’.

Deleting a lexicon note:
Click footnote link in notes field.
Click ‘Edit’.
Click ‘Delete Note’ and confirm.

Etym
Accessible via etymon view

Adding an etymon note:
Click ‘Add a note’ underneath etymon heading in etymon view.
Type note in text box, using markup guidelines for special formatting, (Note type will be ‘Internal’.)
Click ‘Add Note’.
The new note will appear under any pre-existing etymon notes.
Editing an etymon note:
Locate relevant note underneath heading in etymon view.
Make desired edits in the text box on the right. Markup guidelines can be toggled on & off by clicking ‘markup’.
Click ‘Save Note’.

Deleting an etymon note:
Locate relevant note underneath heading in etymon view.
Click ‘Delete Note’ button on the right and confirm.

Etymon Subgroup
Accessible via etymon view

Adding an etymon subgroup note:
Click [+ ] in notes field of the subgroup header (in etymon view).
Type note in text box, using markup guidelines for special formatting. (Note type will be ‘Internal’.)
Click ‘Add Note’.
A footnote reference will appear in the notes field, linking to the note text at the bottom of the page.

Editing an etymon subgroup note:
Click footnote link in notes field of subgroup heading to jump down to note.
Click ‘Edit’ button next to note.
Make desired edits in the text box on the right. Markup guidelines can be toggled on & off by clicking ‘markup’.
Click ‘Save Note’.

Deleting an etymon subgroup note:
Click footnote link in notes field of subgroup heading.
Click ‘Edit’.
Click ‘Delete Note’ and confirm.

Tagging
Associating lexical records with etyma

Use the etyma table to determine the tag number of the relevant etymon.
Locate the lexical item to tag in the lexicon table.

(If applicable) Click in the form field and segment the lexical item into hypothesized morphemes, then press ENTER or TAB:
- Use SPACE to insert a STEDT delimiter (◦), which separates syllables not already
  segmented by a dash, space, period, tone number, etc.
  e.g. azü → a◦zü (a and zü can now be tagged separately)

- Use a VERTICAL BAR (|) placed immediately after an existing delimiter from the
  original source (dash, period, etc.) to override the delimiter and treat the separated
  syllables/morphemes as a unit.
  e.g. a-ji → a|ji (aji can now be tagged as a unit)

Click in the (B) analysis field (the one with your username) and type in the relevant tags, then press
ENTER or TAB:

- Enter these symbols for each type of morpheme:

<table>
<thead>
<tr>
<th>symbol</th>
<th>morpheme type</th>
</tr>
</thead>
<tbody>
<tr>
<td>(tag #)</td>
<td>reflex of etymon with tag #</td>
</tr>
<tr>
<td>p</td>
<td>prefix</td>
</tr>
<tr>
<td>s</td>
<td>suffix</td>
</tr>
<tr>
<td>m</td>
<td>unanalyzed morpheme</td>
</tr>
<tr>
<td>o</td>
<td>onomatopoeia</td>
</tr>
<tr>
<td>c</td>
<td>borrowing from Chinese</td>
</tr>
<tr>
<td>n</td>
<td>borrowing from Nepali</td>
</tr>
<tr>
<td>b</td>
<td>other borrowing</td>
</tr>
</tbody>
</table>

- Separate the tags corresponding to each morpheme by commas. Examples:

  1) Mongsen Ao (Central Naga) ‘dog meat’ (lexicon rn 450685) is tagged as:

      analysis      form
      p,1764,34    [a]-ji-sà?

      [a]    prefix
      ji    reflex of #1764 PTB *d-kʷəy-n DOG
      sà?   reflex of #34 PTB *sy-a-n FLESH / MEAT / GAME ANIMAL

  2) Mizo (Central Chin) ‘forty’ (lexicon rn 258320) is tagged as:

      analysis      form
      4678,2409    såwm-li

      såwm     reflex of #4678 PKC *soom TEN (no higher PTB etymon yet)
      li      reflex of #2409 PTB *b-ła-y FOUR

Your tagging now appears in the (B) analysis field. The morphemes in the form field are also
converted to hyperlinks, which show etymon information in a popup when the cursor is hovered.
over them. If your tagging differs from the STEDT tagging, the conflicting morphemes are highlighted in magenta.

Tagging can be edited by clicking in the \textit{(B) analysis} field.

If STEDT decides to adopt your analysis, your tagging will be ‘approved’ and transferred to the \textit{stedt’s analysis} field. However, if you don’t want it approved yet, you can append a question mark (? ) to the end of any tag number to mark your analysis as provisional.

**Creating Etyma**

\textit{Entering your own reconstructions}

Open the etyma table.

Click the ‘Add a record’ link at the bottom.

\textit{(Optional but highly encouraged)} Use the glosswords table to determine the \textit{semkey} (semantic category) for your proto-gloss. Enter the \textit{semkey} in the \textit{chapter} box.

Enter the proto-form and proto-gloss.

Choose the proto-language in the \textit{grpid} box.

Enter any notes in the \textit{notes} box.

Click ‘Add Record.’ Your new etymon appears in the list above, and the tag number can now be used for tagging lexicon records.

The etymon fields can be edited by clicking on them, making changes, and pressing ENTER or TAB.

**Adding a Meso-reconstruction to an Etymon**

\textit{Filling in the layers of analysis}

Open the etymon view of a particular reconstruction by clicking on its tag number in the etyma table or the lexicon table (in the etymon popup).

Scroll down to the subgroup for which you want to create a meso-reconstruction. Note that only genetic groups allow meso-reconstructions (e.g. Kuki-Chin but not “Naga”).

Click ‘add/edit reconstruction’ in the subgroup heading.

Enter the proto-form and proto-gloss in the blank fields and click ‘Save Reconstructions’. Your new meso-reconstruction appears in the subgroup heading.

Meso-reconstructions owned by you can be edited/deleted by clicking the ‘add/edit reconstruction’ link and making changes. (To delete a meso-reconstruction, click the ‘delete’ checkbox.) Then click ‘Save Reconstructions’.