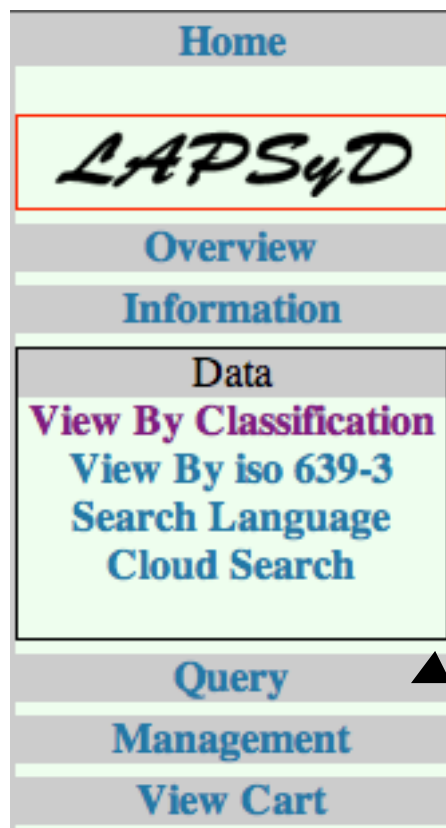


LAPSyD queries

LAPSyD includes powerful tools to build queries based on the presence or absence of segments, features or prosodic properties and their distribution. Queries can be constructed on all types of categorical data in LAPSyD. Terms used in the commentary fields can be searched for using “Cloud search”. Results can be saved and exported or visualized in graphs and maps. All queries are over the set of languages in the database. This documentation provides a guide to constructing queries, but please note that LAPSyD is an ongoing project and documentation is likely to be always a little bit behind the ongoing development.

To see query options click on “Query” in main menu on left of page after “browse data” or log-in. There are three options for query types



Query options

Search by symbol from table

Build query using Boolean operators

Search for contrast

The first option “symbol from table” allows a search for a limited set of specific segments selected from a table of consonant or vowel symbols.

Clicking on this item brings up the explanation shown in the box below. The result is a list of the languages (names and codes) that satisfy the specified search criteria. Clicking on a language name or code opens a new window displaying the LAPSyD entry for that language. Other options for visualizing the results, such as on a map, are shown below the list. Check the boxes for what you want to see, then click on “Show visualizations.”

Query from phonetic chart

This search tool allows you to find the languages in LAPSyD which either have or lack specific segments in their phonological inventory.

Click on one or more of the symbols in the consonant and vowel charts below. To include a particular segment in the search, click once on the appropriate IPA symbol. It will turn blue, like this: **t**. To exclude a particular segment from the search, click on the appropriate IPA symbol twice, so that it turns red, like this: **t**. Clicking on a symbol a third time will deselect it entirely. You can also click on the RESET label in the lower right hand corner of the table to deselect all symbols.

The set of languages that matches the inventory that you have specified is listed below the table. This set consists of the languages whose phonological inventories possess any segments selected as present (in blue) and lack the ones selected as absent (in red).

Only a small subset of possible segments is shown in these charts. Searches for other segments must use a different query option.

Example: clicking once on /b/ and twice on /p/ will provide the list of languages that have a /b/ phoneme but no /p/ (the “missing /p/” pattern). Note that the resulting list will not include languages where /p^h/ but not /p/ occurs and will include languages that have only a voiced plosive series. To find languages with **either** /p/ **or** /p^h/ you can use the Boolean search option in the Query menu, or run separate searches here and add the results together. To eliminate languages with no voiceless plosives /k/ could be added to the search

Boolean Query

Using constraints to build queries in LAPSyD

Query Languages in : All Selected

priority	constraint	sort	action
1	Unicode constraint. <input type="checkbox"/> Strict <input type="button" value="show matches"/> <input type="button" value="Include"/> <input type="button" value="Exclude"/> <input type="button" value="OR"/> <input type="button" value="AND"/>	<input type="button" value="up"/> <input type="button" value="down"/>	<input type="button" value="filter"/>

Add constraint(s)
select type


At the time these instructions were written the initial display for building queries looks like the picture above. The default setting is for a Unicode constraint defined over vowels and all languages. A pull-down menu under “Add constraint(s)” allows for selection of other types:

Add constraint(s)

- select type
- Unicode
- Feature
- Family
- Area
- Complements
- Syllabic index

Unicode Constraint

A **Unicode** constraint selects languages including the chosen symbol(s). Other options in the “Add constraints” menu will be described later.

- Click on any symbol in the vowel table to add it to the search
- Click on “To Consonant” to bring up a table of consonant symbols
- Click on any symbol in the consonant table to add it to the search
- Symbols selected will be displayed in the constraint definition window
- To add diacritics or other modifications to a symbol click on it in the constraint definition window
- Select whether the search is for languages which **INCLUDE** or **EXCLUDE** the symbols
- If more than one symbol is chosen, select whether the operator between them is **AND** or **OR**
- Check the box “strict” by clicking on it to define a search for just the symbol(s) chosen, excluding all combinations in which the symbol(s) occur together with some other element. Leave this box unchecked to search for all occurrences of the symbol(s) including those in combinations.
- Click on  to run the search.

The example below shows symbols **i** and **y** selected with the options “Include” and “AND” and the box “strict” checked. This search finds the languages that contain both /y/ and /i/ in their inventories, but does not look for the occurrence of, for example, diphthong /ai/ or nasalized /ỹ/.

priority	constraint	sort	action
1	Unicode constraint. <input checked="" type="checkbox"/> Strict <input type="button" value="show matches"/> <input type="button" value="Include"/> <input type="button" value="Exclude"/> <input type="button" value="OR"/> <input type="button" value="AND"/> <input type="text" value="i"/> <input type="text" value="y"/>	<input type="button" value="up"/> <input type="button" value="down"/>	<input type="button" value="filter"/>

The search finds 20 languages (among the 286 languages public as of January 15 2014):

order	constraint(s)	filter(s)	sort
1	Strict Unicode constraint. include all these iy	20 languages <input type="button" value="filter"/>	<input type="button" value="up"/> <input type="button" value="down"/>



A more complex example is shown below.

- A first Unicode constraint is defined to exclude /p/ strictly defined
- A second Unicode constraint is added (select “type” = Unicode, click “add”)
- The second constraint looks for languages which include /mb/ strictly defined (click on symbol **m** in consonant table, then “add consonant symbol” and click on **b**)

- click on to run the search

prior	constraint	sort
1	Unicode constraint. <input checked="" type="checkbox"/> Strict <div style="background-color: yellow; padding: 5px;"> Include Exclude OR AND </div> <p>p</p>	<input type="button" value="up"/> <input checked="" type="button" value="x"/> <input type="button" value="down"/>
2	Unicode constraint. <input checked="" type="checkbox"/> Strict <div style="background-color: yellow; padding: 5px;"> Include Exclude OR AND </div> <p>mb</p>	<input type="button" value="up"/> <input checked="" type="button" value="x"/> <input type="button" value="down"/>

The result is as shown below:

order	constraint(s)	filter(s)	sort
1	Strict Unicode constraint. exclude all these :p	32 languages 	<input type="button" value="▲ up"/> <input checked="" type="button" value="x"/> <input type="button" value="▼ down"/>
2	Strict Unicode constraint. include all these :mb	5 languages 	<input type="button" value="▲ up"/> <input checked="" type="button" value="x"/> <input type="button" value="▼ down"/>

The first constraint finds 32 languages with no /p/ phoneme. Within this selection, 5 possess a bilabial prenasalized voiced plosive /mb/. The first constraint can be seen as a filter, determining which languages are submitted to the second. The order of application of the constraints can be changed by using the “up” and “down” buttons in the sort column on the right.

A constraint can be deleted by clicking on the red-highlighted **X** in the sort column.

Feature constraint

Searches can be performed using the **feature** set which defines segments in LAPSyD. Users should become familiar with the feature set by reading the description under the “Information” menu item.

- select “Feature” in “select type” pull-down menu
- click on “Add”: a new line will be added to the constraint definition window
- click on red-highlighted X in the sort column to remove default Unicode constraint line
- select the major segment class desired in the “segment type” menu: Consonants, Vowels, Diphthongs, Clicks (segregation of clicks from other consonants is a pragmatic not a theoretical choice)

These steps are described in some more detail in the next sections.

priority	constraint	sort	action
1	Unicode constraint. <input type="checkbox"/> Strict show matches <div style="border: 1px solid gray; padding: 2px; display: inline-block;"> Include Exclude OR AND </div>	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="display: flex; align-items: center; margin-bottom: 5px;"> ▲ up ✕ </div> <div style="display: flex; align-items: center;"> ▼ down </div> </div>	
2	Feature count constraint. show matches	<div style="display: flex; flex-direction: column; align-items: center;"> <div style="display: flex; align-items: center; margin-bottom: 5px;"> ▲ up ✕ </div> <div style="display: flex; align-items: center;"> ▼ down </div> </div>	

Segment type
 consonant
 ✓ vowel
 diphthong
 click

height(s)

+

backness(s)

+

Lip(s)

+

Class

+

>
 =
 <

0



+

-

✓
Set constraint

For convenience features are grouped into four sets, which depend partly on the segment type chosen as shown in the table below:

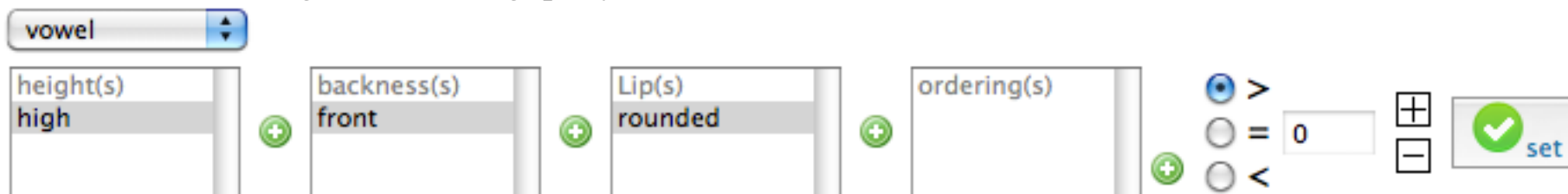
	Group 1	Group 2	Group 3	Group 4
Vowel	Height & Nasalization	Backness, Secondary articulation	Voicing, Timing, Lips	Custom features
Consonant	Manner (Stricture, Airstream, Release, etc)	Place, Secondary articulation	Voicing, Timing,	Custom features
Diphthong	Height & Nasalization	Backness, Secondary articulation	Voicing, Timing, Lips	Ordering, Custom features
Click				

- Click on the green-highlighted plus (+) to see the individual features in each group
- Select features for your search by checking the small white box preceding the feature name
- Click on  to confirm your selections from each group; selected features will appear in the box for the feature group
- When finished with feature selection set numeric limits using the radio buttons to the right
- Limits can be set equal to (=) a given number, or relative to an upper or lower limit using the “more than” (>) or “less than” (<) settings
- Default setting is for “more than zero” occurrences (i.e. search will find all occurrences of the feature)
- To change from the default click on the appropriate radio button (=, >, <), and then use plus (+) or minus (-) to adjust the limit
- Finally, click on  at the right of the boxes to set the search parameters

- To find occurrences between a higher and lower limit use two feature searches, one with upper limit and one with lower limit.

A Feature Query example:

Defining the following query



..... generates this line in the constraint definition window, which enables the user to check if the query corresponds to the desired one.

prior	constraint	sort						
1	<p>Feature count constraint.</p> <table border="1"> <thead> <tr> <th>Type</th> <th>Dimension</th> <th>Criterion</th> </tr> </thead> <tbody> <tr> <td>vowel</td> <td>high front rounded</td> <td>more than 0</td> </tr> </tbody> </table>	Type	Dimension	Criterion	vowel	high front rounded	more than 0	<p>up <input type="checkbox"/></p> <p>down <input type="checkbox"/></p>
Type	Dimension	Criterion						
vowel	high front rounded	more than 0						

This query will find all languages with at least one front rounded vowel (20 among 286 “public” languages at time of writing), and make counts of how many such segments occur in each (hence “feature *count*” constraint).

Visualizing output from feature query

Results of a feature query (as well as other types) can be visualized in various ways, saved or exported. Choice of query type and of visualization selected both affect how results are displayed. Output types include:

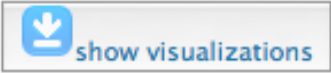
- lists
- tables
- maps
- histograms

Visualization choices are shown in the panel below the constraint definition window. A list of languages corresponding to the search criteria is selected as default. Click on “show visualizations” to see output.

Select visualizations :

- Family distribution.
- Map distribution.
- Show Language List.
- Aggregate Vowel inventory.
- Aggregate Consonant inventory.
- Co-occurrence frequency.
- Canonical form information.

Languages selected.

 show visualizations

The visualization options in this list generate the following outputs:

Family distribution: histogram showing number of found languages by language family. Mouseover shows name and number of languages

Map distribution: location of found languages on world map; colors show areal/genetic groups

Show Language list: list of found languages with name, ISO code, family affiliation

Aggregate Vowel inventory: table(s) showing all vowels (and diphthongs, if any) in the found languages, with number of languages in the set in which each vowel occurs

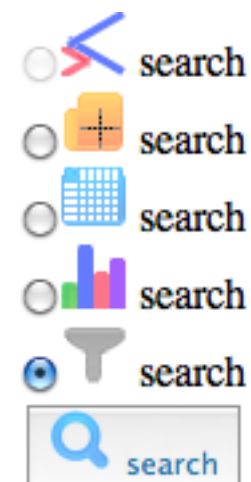
Aggregate Consonant inventory: table showing all consonants in the found languages, with number of languages in the set in which each consonant occurs

Co-occurrence frequency: table showing co-occurrence of all segments in the inventories of the found languages. Each row represents a given segment; rows are ordered from highest to lowest number of languages in which they appear. Each column represents a language, identified by its ISO code. An X signifies that the given segment occurs in the given language.





Canonical form information: table showing syllable structure information by language.


Tables, histograms, comparators

Depending on how a query is framed, the output might be best viewed as a table, a histogram, a contingency table, or a comparison set. These options are selected in the options presented below the “Add constraint” pull-down menu.




The example below shows two feature constraints defined looking for languages which have (1) any voiced stops and/or (2) any voiced fricatives:



priority	constraint	sort	action						
1	<p>Feature count constraint. show matches</p> <table border="1"> <tr> <td>Type</td> <td>Dimension</td> <td>Criterion</td> </tr> <tr> <td>consonant</td> <td>stop voiced</td> <td>more than 0</td> </tr> </table>	Type	Dimension	Criterion	consonant	stop voiced	more than 0	<p>up </p> <p>down</p>	
Type	Dimension	Criterion							
consonant	stop voiced	more than 0							
2	<p>Feature count constraint. show matches</p> <table border="1"> <tr> <td>Type</td> <td>Dimension</td> <td>Criterion</td> </tr> <tr> <td>consonant</td> <td>fricative voiced</td> <td>more than 0</td> </tr> </table>	Type	Dimension	Criterion	consonant	fricative voiced	more than 0	<p>up </p> <p>down</p>	
Type	Dimension	Criterion							
consonant	fricative voiced	more than 0							

Clicking on the icon for a contingency table  search produces the output below. Each of the constraints is independently evaluated and the intersection of the results is displayed. Result of applying the first constraint is shown in columns, and of the second in rows. In this case, of the 286 languages examined, 200 have one or more voiced stops. Of these, 96 also have one or more voiced fricatives, and 104 have none. 125 of the languages have voiced fricatives, and 161 have none.

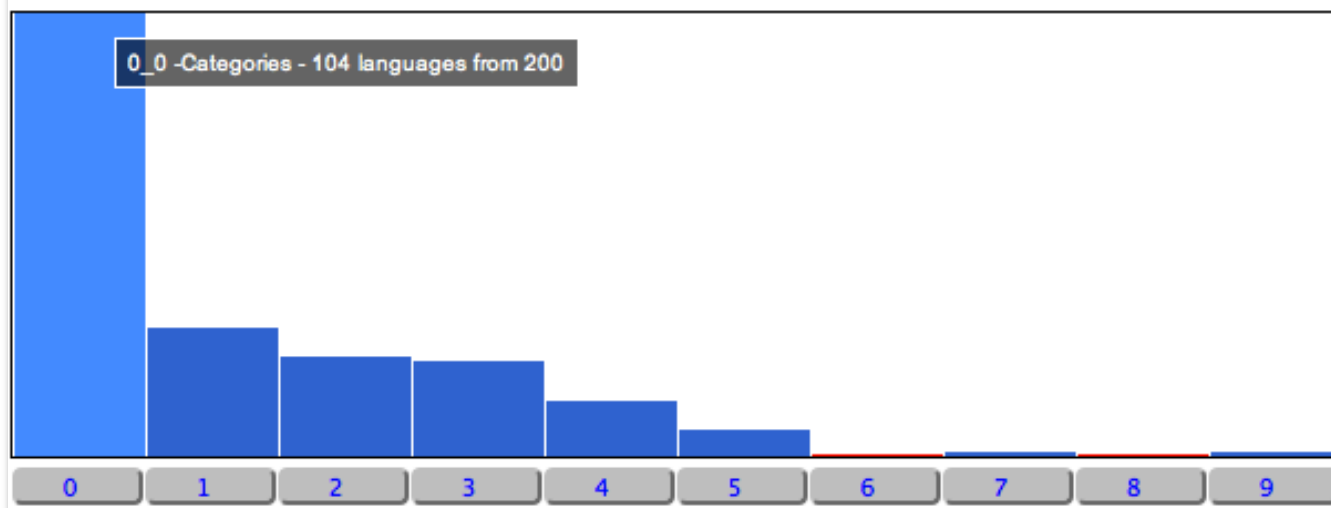
order	constraint(s)	filter(s)	sort									
1	Feature count constraint. <table border="1"> <thead> <tr> <th>Type</th> <th>Dimension</th> <th>Criterion</th> </tr> </thead> <tbody> <tr> <td>consonant</td> <td>stop</td> <td>more than 0</td> </tr> <tr> <td></td> <td>voiced</td> <td></td> </tr> </tbody> </table>	Type	Dimension	Criterion	consonant	stop	more than 0		voiced			<input type="button" value="up"/> <input type="button" value="down"/>
Type	Dimension	Criterion										
consonant	stop	more than 0										
	voiced											
2	Feature count constraint. <table border="1"> <thead> <tr> <th>Type</th> <th>Dimension</th> <th>Criterion</th> </tr> </thead> <tbody> <tr> <td>consonant</td> <td>fricative</td> <td>more than 0</td> </tr> <tr> <td></td> <td>voiced</td> <td></td> </tr> </tbody> </table>	Type	Dimension	Criterion	consonant	fricative	more than 0		voiced			<input type="button" value="up"/> <input type="button" value="down"/>
Type	Dimension	Criterion										
consonant	fricative	more than 0										
	voiced											

286 Language(s)			
	YES	NO	1. count
YES	96	104	200
NO	29	57	86
2. count	125	161	286

If instead, a histogram is requested by clicking on the button  search, the result will be as shown below. In this case, the first constraint acts as a filter. This selects the 200 languages which have voiced stops. The histogram then shows how many voiced fricatives occur in these languages. The first bar in the histogram (selected) shows how many of these have 0 voiced fricatives (104 — as we know from the contingency table).

order	constraint(s)	filter(s)	sort						
1	Feature count constraint. <table border="1"> <tr> <td>Type</td> <td>Dimension</td> <td>Criterion</td> </tr> <tr> <td>consonant</td> <td>stop voiced</td> <td>more than 0</td> </tr> </table>	Type	Dimension	Criterion	consonant	stop voiced	more than 0	200 languages 	<input type="button" value="up"/> <input type="button" value="down"/>
Type	Dimension	Criterion							
consonant	stop voiced	more than 0							
2	Feature count constraint. <table border="1"> <tr> <td>Type</td> <td>Dimension</td> <td>Criterion</td> </tr> <tr> <td>consonant</td> <td>fricative voiced</td> <td>more than 0</td> </tr> </table>	Type	Dimension	Criterion	consonant	fricative voiced	more than 0		<input type="button" value="up"/> <input type="button" value="down"/>
Type	Dimension	Criterion							
consonant	fricative voiced	more than 0							

200 Language(s)



In these and other cases, the icons on the right can be used to export data to Excel for further analysis or graphing, to plot the data according to the groups established by the search criteria, or to show the split according to the six established groups.



Export data in Excel format

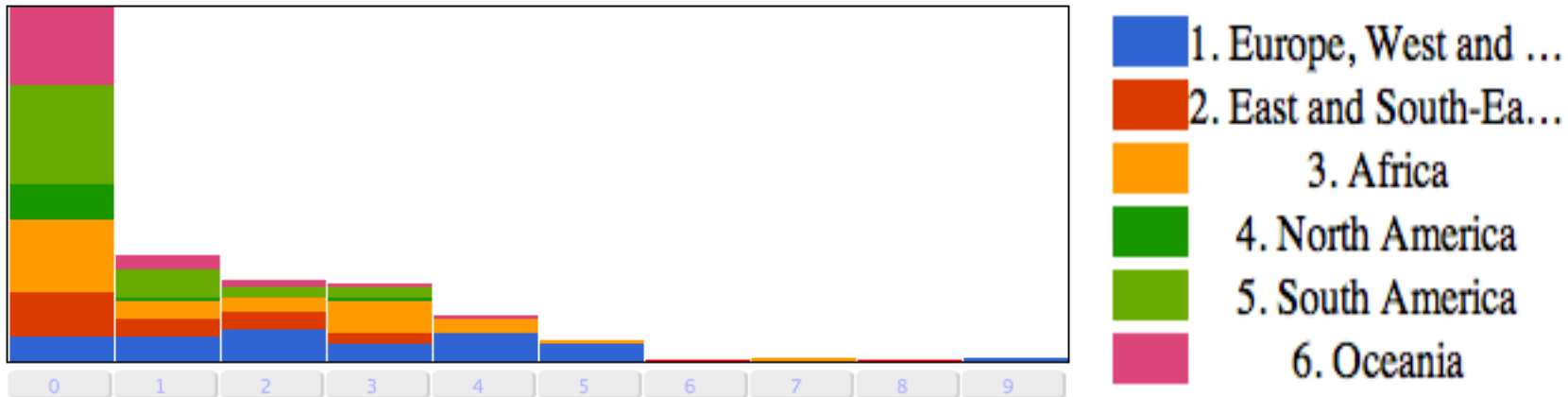


Show groups defined by search criteria

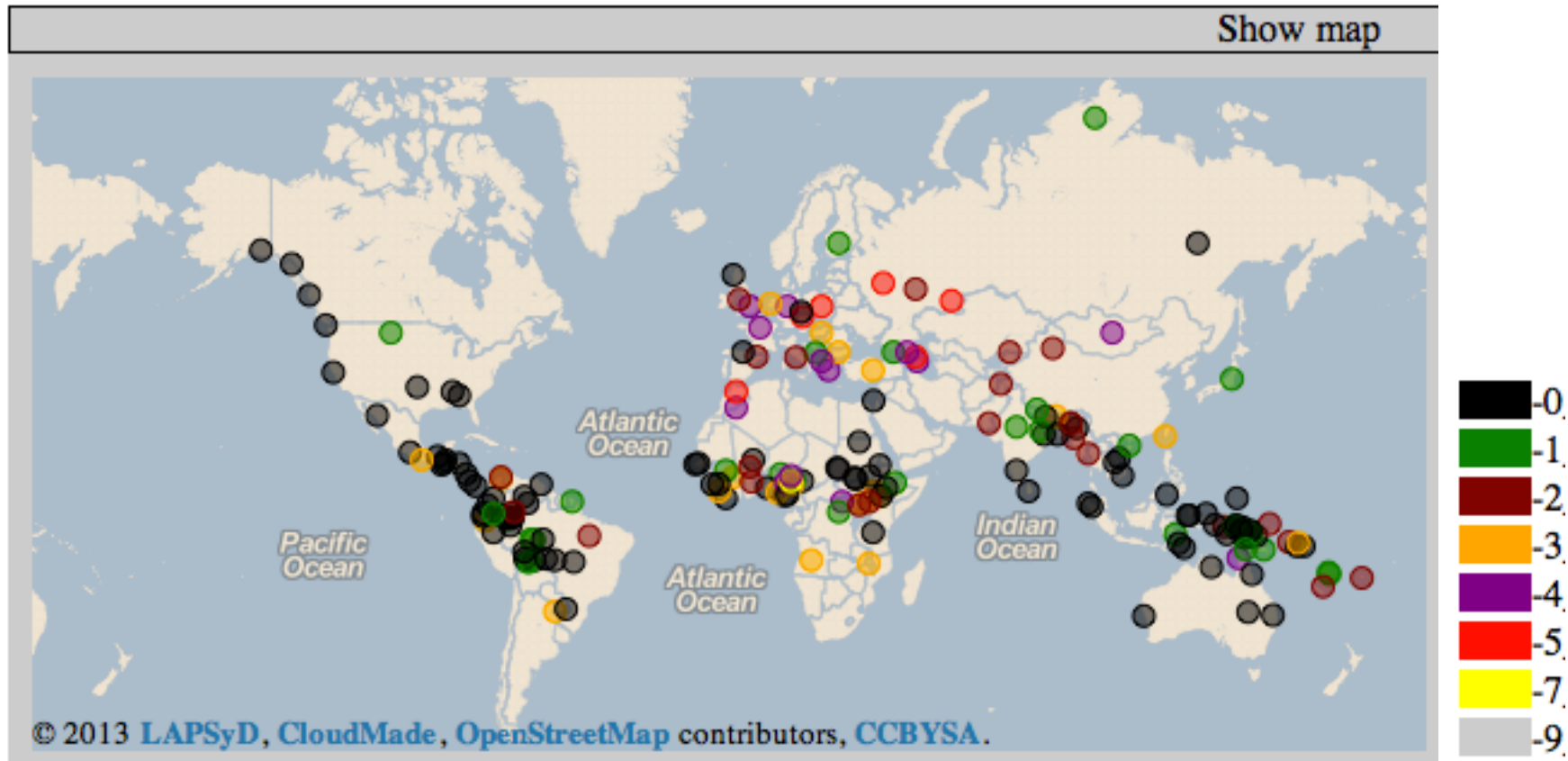


Show areal split

The areal split can be shown in the histogram, as below:



The division into groups can be plotted on a map, as below, where black dots show location of languages with voiced stops but zero voiced fricatives, and so on according to the key:



The maps produced at this or other steps can be viewed on a larger scale by clicking on the display. The center of the map can be altered by dragging, and the map can be saved via a screen dump.