

# Speak For Yersel survey setup instructions

## 1. Introduction

Speak For Yersel is a geospatial linguistic survey tool that enables internet users from a designated geographical area to answer multiple choice questions, with answers then being incorporated into interactive maps in real-time. The tool provides a setup facility through which one or more morphology, lexis and phonology surveys featuring any number of questions can be created, plus facilities for exporting data from the tool once it is in use. This document provides step-by-step instructions to creating the data structures the tool requires, the setup facility, customisation and exporting data.

## 2. The data

The tool requires geospatial data about the geographical area of study and survey questions with associated multiple-choice answers. Survey questions may also include images used to illustrate the questions and phonological survey questions may also include sound files. Data should be prepared prior to beginning the installation of the tool.

A complete set of sample data for Wales is included in the 'example' directory. Installing a test version of the tool using this sample data is a helpful way to understand both the tool and the data structures. Instructions for installing the tool can be found in Section 3 below.

### 2.1 Geospatial data

Geospatial data requires three hierarchical levels of data:

1. **Regions:** These are the broadest geographical areas and they appear on the interactive maps, dividing the overall area of study into a limited number of broad areas, for example 'Borders' in Scotland. Depending on the area of study there should ideally be no more than 10-20 regions and users will be able to highlight a region on the maps, which displays the region name and makes it easy to identify answer markers located within the region.  
Region data must be stored as polygons within a single geojson file named 'regions.json', which must be placed in the tool's 'json' directory. Each polygon must have an ID and a label. The names of these fields as found in the geojson file must be specified in the tool's config.php file located in the 'incs' directory. Update the value of the `$siteOptions["regionID"]` and `$siteOptions["regionLabel"]` variables in this file to match the relevant field names in the geojson file. See below for more details about the config file.
2. **Areas:** These are more specific geographical areas within a region, such as postcode areas. Areas must also be stored as polygons within a single geojson file which does not need to be given a specific filename. The area geojson file will be imported into the system during the setup process as described below. Each region may have any number of associated areas, for example 'Borders' in Scotland features 19 areas. Areas are not displayed on any interactive maps but are used as boundaries when generating a random location for a survey answer.
3. **Settlements:** These are the actual named settlements located within an area and an area may have one or more associated settlements. Settlements must be stored in a

CSV file that is uploaded during the setup process as described below. In addition to the settlement names the CSV also includes fields to connect the settlements to areas and areas to regions. Settlements are employed when a user registers with the survey. During this process the user must state where they lived up to the age of 18 and as they begin to type, the settlements that match the letters will appear in a drop-down list. Once a settlement is selected and the registration process is completed the user is then associated with the settlement, its area and region and the system is then able to generate a random location within the user's area each time a survey question is answered.

The CSV should be saved as UTF-8 text with the following columns. Headings must appear in the first row:

- A. Area: the name of the area, which must exactly match the area name as it appears in the areas geojson file
- B. Region: the name of the region, which must exactly match the region name as it appears in regions.json
- C. Additional group 1: if areas have multiple groupings in addition to region one can be added here. This column must appear, even if it is blank
- D. Additional group 2: as above
- E. Columns from E onwards may contain settlement names (one per column with as many columns as required). Settlements should either appear on their own (e.g. 'Aberfeldy') or with a county name, separated from the settlement name with a comma (e.g. 'Aberfeldy, Perth and Kinross')

## 2.2 Survey data

The tool features three survey types: morphology, lexis and phonology. Multiple surveys of each type can be set up and not all survey types need to be present. Surveys are created and populated with questions and answer options during the setup process as described below. Each survey's data is imported during this process via CSVs, which should be saved as UTF-8 text featuring headings in the first row. The structures of each survey type CSV and a description of the column contents appears below. Note that where a column is optional it must still appear in the CSV, even if it contains no data.

### *Morphology*

- A. ID: a unique identifier for the question. This must be unique across all surveys and should only consist of letters, numbers, dashes and underscores, for example 'morphology-001'.
- B. Order: the order the question will appear in. This must be an integer and mainly exists in the event that questions need to be reordered directly through the database at a later date.
- C. Variable: An optional technical description of the feature. Is not publicly visible but may be of use to researchers when exporting the data
- D. Question: The question that will be presented to users in the survey. May include HTML, such as **bold** or *italic* text
- E. Map name: The text that will appear as a button to identify the question in the list of maps

- F. Max answers: an integer defining how many answer options a user can select. '1' should be the default.
- G. Picture filename: the filename (including extension) of a picture if the question should have an accompanying image. Optional. Image files should be placed in the 'media' directory of the tool. Be aware that filenames and their file extensions are case sensitive.
- H. Picture alt text: Descriptive text that is associated with the image, for example 'An ice cream cone'. Only visible to screen readers or when the image fails to load. Optional
- I. Picture credit: A credit for the image, if required. Will be displayed underneath the image and may include HTML (e.g. a link to the image owner's website). Optional
- J. Xrefs: The IDs of one or more questions in the same survey or another survey, with multiple IDs separated by a bar character (|). Not publicly visible but may be useful for researchers. Optional
- K. Notes: Notes relating to the question for researchers to reference. Not publicly visible. Optional
- L. Columns L onwards should include the answer options with one per column. Will appear as the text of the answer option buttons in the survey.

### Lexis

- A. ID: a unique identifier for the question. This must be unique across all surveys and should only consist of letters, numbers, dashes and underscores, for example 'lexis-001'.
- B. Order: the order the question will appear in. This must be an integer and mainly exists in the event that questions need to be reordered directly through the database at a later date.
- C. Variable: An optional technical description of the feature. Is not publicly visible but may be of use to researchers when exporting the data
- D. Question: The question that will be presented to users in the survey. May include HTML, such as **bold** or *italic* text
- E. Map name: The text that will appear as a button to identify the question in the list of maps
- F. Max answers: an integer defining how many answer options a user can select. '1' should be the default.
- G. Picture filename: the filename (including extension) of a picture if the question should have an accompanying image. Optional. Image files should be placed in the 'media' directory of the tool. Be aware that filenames and their file extensions are case sensitive.
- H. Picture alt text: Descriptive text that is associated with the image, for example 'An ice cream cone'. Only visible to screen readers or when the image fails to load. Optional
- I. Picture credit: A credit for the image, if required. Will be displayed underneath the image and may include HTML (e.g. a link to the image owner's website). Optional
- J. Xrefs: The IDs of one or more questions in the same survey or another survey, with multiple IDs separated by a bar character (|). Not publicly visible but may be useful for researchers. Optional
- K. Notes: Notes relating to the question for researchers to reference. Not publicly visible. Optional

- L. Part of speech: The part of speech of the word associated with the question. Not publicly visible. Optional
- M. Change: a description of the associated linguistic change. Not publicly visible. Optional
- N. Columns N onwards should include the answer options with one per column. Will appear as the text of the answer option buttons in the survey.

### *Phonology*

- A. ID: a unique identifier for the question. This must be unique across all surveys and should only consist of letters, numbers, dashes and underscores, for example 'phonology-001'.
- B. Order: the order the question will appear in. This must be an integer and mainly exists in the event that questions need to be reordered directly through the database at a later date.
- C. Variable: An optional technical description of the feature. Is not publicly visible but may be of use to researchers when exporting the data
- D. Question: The question that will be presented to users in the survey. May include HTML, such as **<b>bold</b>** or *<i>italic</i>* text
- E. Map name: The text that will appear as a button to identify the question in the list of maps
- F. Max answers: an integer defining how many answer options a user can select. '1' should be the default.
- G. Picture filename: the filename (including extension) of a picture if the question should have an accompanying image. Optional. Image files should be placed in the 'media' directory of the tool. Be aware that filenames and their file extensions are case sensitive.
- H. Picture alt text: Descriptive text that is associated with the image, for example 'An ice cream cone'. Only visible to screen readers or when the image fails to load. Optional
- I. Picture credit: A credit for the image, if required. Will be displayed underneath the image and may include HTML (e.g. a link to the image owner's website). Optional
- J. Xrefs: The IDs of one or more questions in the same survey or another survey, with multiple IDs separated by a bar character (|). Not publicly visible but may be useful for researchers. Optional
- K. Notes: Notes relating to the question for researchers to reference. Not publicly visible. Optional
- L. Audio filename stem: If the question has an associated set of audio files (one per answer option) these must have a standardised filename consisting of a single stem followed by a dash followed by the number of the answer option. Enter the stem in this column. For example, if the question relates to the pronunciation of 'bat' and there are four answer options / audio files then the stem may be 'bat'. The audio files (which must be MP3) should then be given the filenames 'bat-1.mp3', 'bat-2.mp3', 'bat-3.mp3' and 'bat-4.mp3' and placed in the tool's 'media' directory. Be aware that filenames and their file extensions are case sensitive.
- M. Columns M onwards should include the answer options with one per column. Will appear as the text of the answer option buttons in the survey unless an audio stem is given, in which case a 'play' icon will appear on each button and the text provided will not be publicly visible, but will appear in the exported data.

## 3. Setting up the tool

### 3.1 Technical requirements

The tool requires a web server, the PHP scripting language and a relational database such as MySQL. It uses the Bootstrap frontend toolkit, the jQuery JavaScript library, the Leaflet JavaScript mapping library and the Turf JavaScript geospatial analysis library.

### 3.2 Initial setup

The tool's files and directories can be placed in the web-accessible root directory of a website on your server or within a subdirectory. A database and associated user must also be created for the tool. The user must have privileges to create and alter database tables for the database. Database credentials (host, database name, database user and password) must be added to the db.php file located in the tool's 'incs' directory.

### 3.3 Configuration options

Options for your instance of the tool should be added to the 'config.php' file located in the tool's 'incs' directory prior to beginning setup. These options can be updated at a later date, but updating some options (e.g. the database prefix) may cause the tool to stop working.

The following options are included:

- `$siteOptions["siteTitle"]`: The name of the site, which will appear in the heading of every page
- `$siteOptions["surveyArea"]`: The area the survey covers. Appears in various places throughout the site, such as the registration page
- `$siteOptions["launchDate"]`: Used on the 'project staff' page when limiting statistics and data exports. Must use the format "yyyy-mm-dd"
- `$siteOptions["siteLogo"]`: The filename (including extension) of the logo used in the site header (if applicable). The corresponding image file should be added to the tool's 'graphics' directory.
- `$siteOptions["dbPrefix"]`: An optional prefix that will be added to all database tables the tool creates (e.g. 'roi\_'). Useful if multiple surveys are stored in one database. Should only include alphanumeric characters, dashes and underscores
- `$siteOptions["subDirectory"]`: If the site is being set up somewhere other than the web accessible root directory on the server enter the path here, without a final slash. E.g. 'roi'.
- `$siteOptions["defaultLatLng"]`: The location all maps will centre on by default. Latitude then longitude as floats, separated by a comma. E.g. "54.6088, -6.8852"
- `$siteOptions["defaultZoom"]`: The default zoom (integer or float) at which all maps will load, for example '8'
- `$siteOptions["gTag"]`: If your site uses Google Analytics supply the GA tag ID here
- `$siteOptions["regionID"]`: The field containing each region's unique identifier in your 'regions.json' file (see the 'data' section above)
- `$siteOptions["regionLabel"]`: The field containing each region's label in your 'regions.json' file (as above)

- `$siteOptions["attributionText"]`: The text that will appear in the map's 'attribution' popup. Can contain HTML, but ensure backslashes are used before quotation marks, e.g. `target="_blank"`
- `$siteOptions["otherLanguageQuestion"]`: Set this to 'true' if the survey area features another language and you wish to record speakers' bilingualism
- `$siteOptions["otherLanguage"]`: The name of the other language, for example 'Welsh'.

### 3.4 Other configuration options

#### *Base map*

The original Speak For Yersel used a MapBox base map, which requires a user account and an access token. By default the tool uses a free alternative. You can replace this by opening the 'sfy.js' file located in the tool's 'js' directory in a text editor and updating the 'base map' details that appear near the top of the file.

#### *Marker colours*

Marker colours use a gradient for morphology surveys and a selection of different colours for the other two survey types. To update these open the 'sfy.js' file located in the tool's 'js' directory in a text editor and find the 'ratingColours' and 'lexicalColours' arrays near the top of the file. These can be replaced with any other colour hex codes. Be sure to leave the first array element blank as this is not used.

You will also need to update the corresponding colours in the tool's CSS file, 'sfy.css' located in the tool's 'css' directory. These are named '.rating-' or '.lexical-' with a number corresponding to the array element in the JavaScript file.

#### *Site text*

The text for the site's homepage, 'about' page and 'resources' page is located in the 'index.php' file in the root directory of the tool. The relevant sections are clearly labelled and HTML text can be added to the appropriate sections as required.

### 3.5 Running the setup script

#### *Stage 1*

Once you have created the geospatial and survey data, have entered your database details into 'incs/db.php' and have created the required options in the 'incs/config.php' you are ready to bring everything together and create a survey instance.

Open the file 'setup/index.php' in a text editor and delete the characters '/\*' from the line above 'require("../incs/db.php");'. Save the file and ensure all of the tool's files are uploaded to your web server. Then open the setup script in a web browser: <https://your-domain/subdomain-if-specified/setup/>

You should be presented with a screen like the following:

## Setup Speak For Yersel

### Setup a new Speak For Yersel instance - Stage 1

If you haven't done so already please ensure you have created a database for the project and have added the database name, username and password to the file `incs/db.php`.

Also ensure you have supplied any required details in the file `incs/config.php`. The following details are currently saved:

- **siteTitle:** Speak For Yersel: Northern Ireland
- **surveyArea:** Northern Ireland
- **launchDate:** 2024-03-01
- **siteLogo:**
- **dbPrefix:** ni\_
- **subDirectory:** ni
- **defaultLatLng:** 54.6088, -6.8852
- **defaultZoom:** 8
- **gTag:** [REDACTED]
- **regionID:** OBJECTID
- **regionLabel:** LGDNAME
- **attributionText:** <h3>Linguistic data</h3><p>The linguistic data featured in this resource are copyright of the University of Glasgow.</p><h3>Base map data</h3><p>© <a href="https://openstreetmap.org" target="\_blank">OpenStreetMap</a> contributors, <a href="https://creativecommons.org/licenses/by-sa/2.0/" target="\_blank">CC-BY-SA</a>, Imagery © <a href="https://mapbox.com" target="\_blank">Mapbox</a></p>

If you wish to change these values update `incs/config.php` and reload this page.

Continue

Developed at the University of Glasgow

If any of the listed options are incorrect edit the 'incs/config.php' file and then reload the page. Once all is correct press the 'continue' button and the script will create all of the required database tables.

## Stage 2

Once the database tables have been created a page like the one below will load:

## Setup Speak For Yersel

### Setup a new Speak For Yersel instance - Stage 2

Database tables have been successfully created.

#### Create a survey type

Survey type	Survey name	Page URL
<input type="text" value="v"/>	<input type="text"/>	<input type="text"/>

Introductory text

Continue

Developed at the University of Glasgow

You can now create all of the individual surveys that are required. As discussed above, there are three survey types (morphology, lexis and phonology) and any number of individual surveys can be created, for example multiple phonology surveys, or omitting a morphology survey. Using the form in Stage 2 you can create each survey that is required. Select the 'Survey type' and supply a 'Survey name'. This will be displayed in the website and can be whatever text you prefer. For example, the original SFY survey named the 'phonology' survey 'Sounds about right'.

You must also supply a page URL for the survey. This must be unique and must only include alphanumeric characters, dashes and underscores (i.e. no spaces). Introductory text must also be supplied. This will appear on the survey's introduction page. HTML tags can be entered into this box.

Once you have entered all of the required information press the 'Continue' button and your survey will be saved. The form will reload allowing you to create another survey. Once you have created all required surveys press the 'Continue to stage 3' button.

### Stage 3

In stage 3 you will upload the survey questions and answer options for each of your created surveys. Your surveys will be listed on a page like the following:

#### [Setup Speak For Yersel](#)

Setup a new Speak For Yersel instance - Stage 3

Upload survey questions and answer options

Explanatory text here

Survey Type	Survey Name	Questions	
morphology	I would never say that!	<input type="button" value="Browse..."/>	No file selected.
lexis	Give your word	<input type="button" value="Browse..."/>	No file selected.
phonology	Sounds about right	<input type="button" value="Browse..."/>	No file selected.

Upload Questions

Developed at the University of Glasgow

You should have a CSV file containing your questions and answer options for each survey, as discussed above. Use the 'Browse...' option beside each survey to find and attach the relevant CSV file. Once you've attached them all press the 'Upload Questions' button and a summary page should load such as the following:



## [Setup Speak For Yersel](#)

### Setup a new Speak For Yersel instance - Stage 3

#### Upload survey questions and answer options

Explanatory text here

Survey Type	Survey Name	Questions
morphology	I would never say that!	30
lexis	Give your word	35
phonology	Sounds about right	30

Questions have been uploaded for all surveys.

[Continue](#)

Developed at the University of Glasgow

## Stage 4

Upon pressing the 'Continue' button the script will then process the 'regions.json' file (as discussed above) to extract and list the regions. A page such as the following will load:

## [Setup Speak For Yersel](#)

### Setup a new Speak For Yersel instance - Stage 4

Creating regions:

- 1: Antrim and Newtownabbey
- 2: Armagh City, Banbridge and Craigavon
- 3: Belfast
- 4: Causeway Coast and Glens
- 5: Derry City and Strabane
- 6: Fermanagh and Omagh
- 7: Lisburn and Castlereagh
- 8: Mid and East Antrim
- 9: Mid Ulster
- 10: Newry, Mourne and Down
- 11: Ards and North Down

#### Upload Area JSON file

Explanatory text here

Area GeoJSON File

[Browse...](#) No file selected.

Property Field containing area ID

Property field containing area name

[Continue](#)

You can now import your area geoJSON file (as discussed above). To do so use the 'Browse...' option to attach your geoJSON file then enter the names of the fields that contain each area's ID

and name in the geoJSON data. Note that these field names are case sensitive. For example, in our Northern Ireland areas geoJSON file the area ID field is named 'OBJECTID' and the area name field is 'Name' so these would be added to the corresponding boxes in the form.

Press the 'Continue' button and the areas will be extracted from your geoJSON file.

#### *Stage 5*

This stage will list all of the areas that have been created, and will also display any errors encountered when parsing the geoJSON. A page such as the following will be displayed:

Setup a new Speak For Yersel instance - Stage 5

Creating areas:

1: BT6

2: BT8

3: BT10

4: BT5

5: BT54

6: BT7

8: BelfastCentre

9: BT4

10: BT3

11: BT9

12: BT17

13: BT16

14: BT19

15: BT18

16: BT11

17: BT13

18: BT4

19: BT12

20: BT15

21: BT28

22: BT26

23: BT27

24: BT29

25: BT24

26: BT23

27: BT21

28: BT25

29: BT20

30: BT22

31: BT39

32: BT38

33: BT36

34: BT35

35: BT31

36: BT34

37: BT37

38: BT30

39: BT33

40: BT81

41: BT32

42: BT82

43: BT80

44: BT40

45: BT46

46: BT98

47: BT92

48: BT94

49: BT40

50: BT45

51: BT48

52: BT47

53: BT41

54: BT42

55: BT52

56: BT55

57: BT44

58: BT43

59: BT51

60: BT57

61: BT56

62: BT53

63: BT60

64: BT62

65: BT61

66: BT68

67: BT67

68: BT69

70: BT63

71: BT71

72: BT66

73: BT65

74: BT70

75: BT79

76: BT78

77: BT75

78: BT74

79: BT77

80: BT76

113: Waterside

115: Airmagh

119: Newry

123: Dungannon

125: Omagh

126: Ballymacarrett

134: Short Strand

138: Clarawood

146: Woodstock

159: Cregagh

162: Donegall Pass

164: Taughmonagh

168: NvlBelfast

176: Poleglass

178: TheMarket

180: QUB

182: Ballysillan

183: Strabane

184: Lurgan

186: Craigavon

188: Portadown

193: Coleraine

194: Ballymena

195: Lisburn North

199: Lisburn South

200: PortLaferry

201: Newtownards

203: Antrim Coast

Upload Settlement CSV File

Explanatory text here

Settlement CSV File

Browse...

No File selected.

Continue

With the areas created you can now upload the settlements by attaching your settlements CSV file (see above) to the form and pressing ‘Continue’. This will then process your settlements and associate them with areas and regions. A log of all updates will be displayed, including any errors encountered. A small section of such a page can be viewed below:

#### Setup Speak For Yersel

### Setup a new Speak For Yersel instance - Complete

Associating CSV areas, joining to regions and adding settlements:

- Data for Antrim Coast merged with record from GeoJSON and associated with region 'Mid and East Antrim'. Settlements:
  - Carnlough ( County Antrim)
  - Clogh Mills ( County Antrim)
  - Cushendall ( County Antrim)
  - Cushendun ( County Antrim)
  - Glenarm ( County Antrim)
  - Glenariff or Waterfoot ( County Antrim)
  - Knocknacarry ( County Antrim)
  - Carnalbanagh ( County Antrim)
- Data for Armagh merged with record from GeoJSON and associated with region 'Armagh City, Banbridge and Craigavon'. Settlements:
  - Armagh ( County Armagh)
- Data for Ballymacarrett merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - Ballymacarrett ( Belfast)
  - Newtownards Road ( Belfast)
- Data for Ballymena merged with record from GeoJSON and associated with region 'Mid and East Antrim'. Settlements:
  - Ballymena ( County Antrim)
- Data for Ballysillan merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - Ballysillan ( Belfast)
- Data for BelfastCentre merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - City Centre ( Belfast)
- Data for BT10 merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - Finaghy ( Belfast)
- Data for BT11 merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - Turf Lodge ( Belfast)
  - Lenadoon ( Belfast)
  - Andersonstown ( Belfast)
- Data for BT12 merged with record from GeoJSON and associated with region 'Belfast'. Settlements:
  - Ballymurphy ( Belfast)

#### *Post-setup*

The setup process is now complete and you should test the site (registering as a user and completing the surveys) to ensure that all processes completed successfully. If they didn't you should delete all database tables, make any required changes to your data and begin the setup process again.

Once you are satisfied with the survey you should either delete the ‘setup’ directory from your web server or ensure the script is not executable. To achieve the latter simply reinstate the ‘/\*’ characters in the line above ‘require("../incs/db.php");’ and save the file on your web server.

## 4. Viewing stats and exporting data

Statistics about completed survey questions together with facilities to export survey data can be accessed by loading the page ‘project-staff’ in your web browser. A page similar to the following will load:



## Stats and data download facilities

## Period

Note that to get data for a specific day use the same date as 'from' and 'to'.

From

2024-02-01

Use the format yyyy-mm-dd

To

2024-03-28

Use the format yyyy-mm-dd

Update

The number of users listed is the number of users who registered in the selected period. The number of answers listed are the number of answers submitted in the selected period. Be aware that a user may have registered outside of your selected period but may have submitted answers inside your selected period.

## Total users

A total of **9** people have registered within The Republic of Ireland and **3** from elsewhere.

[Download users from within The Republic of Ireland \(CSV\)](#)[Download users from elsewhere \(CSV\)](#)

## Survey answers

**I would never say that! (morphology)**

There have been **33** submitted answers from within The Republic of Ireland and **67** from elsewhere.

[Download answers from within The Republic of Ireland \(CSV\)](#)[Download answers from elsewhere \(CSV\)](#)

Be aware that by default this page is not password protected and you should create your own password protection for it if required, for example setting up an Apache .htpasswd file.