

## Creating a Custom Dictionary to Run in LIWC2015 for Computational Linguistic Analysis

### Part 1: Intro

#### Part 1A: General Structure of a Custom Linguistic Analysis Dictionary

```
%
1 Dimensiona
2 Dimensionb
3 Dimensionc
%
Word 1
Word 1 2 3
Word 1
Word 2
Word 2 3
```

#### Part 1B: General Steps

1. Define the purpose of the custom linguistic analysis dictionary. What in-world phenomena are you trying to observe? What are you trying to learn? Where will you be acquiring texts for the analysis? How will the texts be processed (or “cleaned”) to enable the analysis?
2. What base language are you planning to use? Why? [It is possible to include several (or more) languages in a custom dictionary, but if a built-in LIWC2015 dictionary is used, there are several base languages to choose from. Custom dictionaries in LIWC2015 do not preclude the use of already built-in dictionaries, which enable a wide variety of observations.]
3. Define constructs as comprehensively as possible. Ensure that these are fairly mutually exclusive with just a little overlap (when needed). Name these in a consistent way.
4. Once the list of constructs is complete, you will put these in order, and number them consecutively (in parentheses behind the construct name).
5. Then, identify all related terms to the construct. Decide if you want to use \*wildcard characters attached to base words to capture stemmed words and word variations from one base.
6. Revise and edit the list.
7. Once the draft is ready to run, start to build your text file. Set up the dimensions.
8. Alphabetize the list of related terms. (This helps make the dictionary somewhat human readable.) Make sure that the related construct number is carried forward with the related terms.
9. Create a .dic file using Word. (\*See the next page for more precise directions on how to achieve this.)
10. Run the custom dictionary in LIWC2015 against known text sets to see how they function.
11. Analyze the custom dictionary for its strengths and weaknesses. Revise the custom dictionary for increased clarity and usability, reliability, and validity... (This is a research instrument now!)
12. Name the dictionary. If there are possible acronyms, make sure that they make sense. Whenever you use the dictionary, cite it.
13. Create a user manual for the dictionary if you plan to share it with others. (Documenting the custom dictionary instrument is a part of good researcher practice anyway.)
14. If you want, you can share your custom dictionary with the LIWC2015 community.

15. Revise it as you think of new constructs and construct-related-words, numbers, dates, emoticons, and other text-related contents.

### \*Creating .dic Files

1. Open MS Word.
2. Click "File" tab in the ribbon.
3. Click "Options" at the bottom left.
4. Select "Proofing" in the "Word Options" window.
5. Click the "Custom Dictionaries" button.
6. Indicate a "New" dictionary.
7. Give the new dictionary a name and save it to the correct location with the .dic file format.
8. Open the .doc file in Word and paste the dictionary (with new words on each line) into the file. Save. Load. Run.
9. If you will be making multiple dictionaries, make a few extras with generic names to serve as templates!

### Part 1C: Selected Resources

#### Linguistic Inquiry and Word Count: LIWC2015

By James W. Pennebaker, Roger J. Booth, Ryan L. Boyd, and Martha E. Francis

[https://s3-us-west-2.amazonaws.com/downloads.liwc.net/LIWC2015\\_OperatorManual.pdf](https://s3-us-west-2.amazonaws.com/downloads.liwc.net/LIWC2015_OperatorManual.pdf)

#### The Development and Psychometric Properties of LIWC2015

By James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn, The University of Texas at Austin

[http://liwc.wpengine.com/wp-content/uploads/2015/11/LIWC2015\\_LanguageManual.pdf](http://liwc.wpengine.com/wp-content/uploads/2015/11/LIWC2015_LanguageManual.pdf)

#### LIWC-ing at Texts for Insights from Linguistic Patterns

By Shalin Hai-Jew, Kansas State University

<http://www.slideshare.net/ShalinHaiJew/liwc-ing-at-texts-for-insights-from-linguistic-patterns>

#### See Ya! Creating a Custom Spatial-Based Linguistic Analysis Dictionary from Social Media Data Sets to Explore American Renunciation of Citizenship

By Shalin Hai-Jew, Kansas State University

<http://www.slideshare.net/ShalinHaiJew/see-ya-creating-a-custom-spatialbased-linguistic-analysis-dictionary-from-social-media-data-sets-to-explore-american-renunciation-of-citizenship>

### Part 1D: Presenter Contact Information

Dr. Shalin Hai-Jew  
 Instructional Designer  
 iTAC  
 Kansas State University  
[shalin@k-state.edu](mailto:shalin@k-state.edu)  
 785-532-5262

## Part 2: Live Custom-Dictionary-Creation Activity!

### Part 2A. The Phenomenon / Phenomena You are Studying

- Also, what sorts of hypotheses would you be interested in exploring?
- What are traditional research methods in this area, and will that affect how you use computational linguistic analysis? How you design a custom dictionary?
- How do place and space figure into your topic? How might you represent both “place” and “space” in your work?

### Part 2B. How Language (Written and / or Spoken) Plays into the Phenomenon / Phenomena:

- Does (do) the phenomenon / phenomena you study instantiate in speech? In conversation? In dialogue? In writing? In microblogging messages?
- What are “place” and “space” elements in related words?

### Part 2C. Where You Plan to Find and / or Create Text Sets for Study:

Primary Sourcing	Secondary Sourcing
<input type="checkbox"/> Gray literature <input type="checkbox"/> Interviews <input type="checkbox"/> Focus groups <input type="checkbox"/> In-world observations <input type="checkbox"/> Photos (self-captured) <input type="checkbox"/> Diagrams	<input type="checkbox"/> Published texts <input type="checkbox"/> Speeches, spoken words, or songs (live or recorded, transcribed) <input type="checkbox"/> Social media <input type="checkbox"/> Maps <input type="checkbox"/> Photos (other-captured)




<b>Category 3</b>	
<b>Constructs (Dimensions)</b>	<b>Related Terms</b>

<b>Category 4</b>	
<b>Constructs (Dimensions)</b>	<b>Related Terms</b>

<b>Category 5</b>	
<b>Constructs (Dimensions)</b>	<b>Related Terms</b>

<b>Category 6</b>	
<b>Constructs (Dimensions)</b>	<b>Related Terms</b>

--	--

**Part 2E: Participant Notes, Observations, and Questions:****Part 3: Post Custom-Dictionary-Creation Activity****Part 3A: Efficacy (or Inefficacy) of the Custom Dictionary**

- **If your custom dictionary is successful, what would you see from results of a run of your custom dictionary? Why?**
- **If your custom dictionary is not so successful, what would you see from the results of a run of your custom dictionary? Why?**
- **Are there other ways to test your custom dictionary besides running it in LIWC2015 against a number of text sets? Are there experts you can tap to evaluate the terms and words?**
- **What are ways to improve your custom dictionary? Why?**
- **What are ways to improve references to “place”? What are ways to improve references to “space”?**
- **What are ways to improve your custom dictionary over time? How can you tell if your custom dictionary is dating out?**