Soaring Higher with MACAWS: Integrating iDDL with Other Web Applications to Enhance Language Learning

June 12, 2021

Mariana Centanin Bertho  marianabertho@email.arizona.edu
Aleksey Novikov         alekseynovikov@email.arizona.edu
Adriana Picoral         adrianaps@email.arizona.edu
Bruna Sommer-Farias     fariasbr@msu.edu
Shelley Staples         slstaples@email.arizona.edu
Corpora & Corpus-based Pedagogy

Corpus:
- a collection of texts, spoken or written (plural = corpora)
- representative of a particular domain of language use
- balanced across genres or registers

Learner corpus:
- corpus of learner production, defined broadly (students, L2 writers/speakers, etc.)

Corpus-based pedagogy - use of corpus or corpus findings to:
- inform curriculum
- create materials (e.g., textbooks or teachers' own materials)
- engage students directly in examining corpus data
Corpora & Corpus-based Pedagogy

Data-driven Learning - DDL

Incorporation of **authentic examples** of language use into classroom activities

- **Main goal:** pattern noticing
- **Pedagogical approach:** inductive learning
- **Concordance lines:** a line of text that shows a search term highlighted with its surrounding context
Interactive Data-driven Learning - iDDL

Corpus-based activities built with **interactive concordance lines**.

Can be integrated with several technologies, such as:

- Learning management systems
- Google applications (Docs, Jamboard, Forms, etc.)
- Websites

---

Experimentar e experiência que todos acabam sendo vale a viagem! Espanha é Cananea tem para oferecer a todos. Cananea uma pequena cidade que está e vi um filme americano todos em espanhol! Eu vi Nerve, e foi interessante e bom. Também, todos estamos mais velhos agora. Finalmente, muitas bebidas deliciosas em todos lugares e não é tão caro. Esses eventos acontecem todos os dias, você só precisa encontrá-los. Em Cananea, eu gostava de fazer experiências muito bonita. Todos os tipos de gente tem coisas para fazer. Eu perguntei como são a escola e todos os aspectos de nossas vidas. Também, the small town that doesn't appear on any of the maps, então Tucson é mais turístic
Our Learner Corpora

MACAWS: Multilingual Academic Corpus of Assignments Writing & Speech

CROW: Corpus and Repository of Writing
**Corpus**
- 2,980 samples of college writing and speech, containing 618,541 words
- Texts in Russian & Portuguese, from 2017-2018
- Filter by L1, course, topic, draft, assignment mode
- Search by word or phrase

**Repository (Coming Soon)**
- Downloadable assignment prompts, syllabi, rubrics, lesson plans, and other course material.
  - Filter by material, document format, course, and date
  - Search the full text of resources; download original format

**Resources & Tutorials**
- Ready-to-use corpus-based language learning activities for teachers and language learners:
  - MACAWS-based language learning activities
- Webinars on what MACAWS is and how to use it in the classroom (for teachers):
  - MACAWS Webinars
- Teacher training workshop materials on how to create language learning materials using MACAWS
  - Workshop materials

**What's new?**
- **March 10, 2020:** Teacher Training Sessions -- Fall and Spring 2020
- **February 7, 2020:** American Association of Teachers of Slavic and East European Languages (AATSEEL) in San Diego, California. “MACAWS Russian: Corpus Design and Creation of Usage-Inspired Pedagogical Materials”
- **November 1, 2019:** New Ways to Analyze Variation (NWAV) in Eugene, Oregon. “The acquisition of preposition+article contractions in L3 Portuguese among different L1-speaking learners: A variationist approach”
- **September 1, 2019:** VIII Encontro Mundial sobre o Ensino do Português (EMEP) in Princeton University, Princeton, NJ. “MACAWS and the teaching of writing: Using a Portuguese learner corpus to create pedagogical materials”
## MACAWS Data

<table>
<thead>
<tr>
<th>L2</th>
<th>Mode</th>
<th># of students</th>
<th># of texts</th>
<th># of words</th>
<th>Average word count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russian</td>
<td>Writing</td>
<td>100</td>
<td>765</td>
<td>104,813</td>
<td>137.00</td>
</tr>
<tr>
<td>Russian</td>
<td>Speech</td>
<td>72</td>
<td>269</td>
<td>19,241</td>
<td>71.50</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Writing</td>
<td>255</td>
<td>2,075</td>
<td>536,168</td>
<td>258.40</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>427</strong></td>
<td><strong>3,109</strong></td>
<td><strong>660,222</strong></td>
<td><strong>212.36</strong></td>
</tr>
</tbody>
</table>
Corpus
- 9,759 samples of college writing, containing 8,347,760 words
- Texts from Fall 2014 to Spring 2018
- Filter by country, gender, assignment, date, TOEFL scores
- Search by word, word family, or phrase

What's new?
- November 2: Arizona Teachers of English to Speakers of Other Languages, "Bridging L2 writing and academic vocabulary through corpus-based activities"
- November 15: Symposium on Second Language Writing, "Using a learner corpus and a repository of pedagogical materials for L2 writing research and teaching"
- March 31, 2020: TESOL 2020

Repository
Downloadable assignment prompts, syllabi, rubrics, lesson plans, and other course materials.
- Filter by material, document format, course, and date
- Search the full text of resources; export to markdown

Resources & Tutorials
- Writing Models
  - Compare how other writers use an English idiom
- Lesson Planning
  - Find sample sentences for a lesson on signal phrases
- Research & Analysis
  - Export Crow data for statistical analysis in R
  - Perform machine-driven queries with the API
- Corpus Building
  - Develop a similar corpus using the source code
Demonstration Outline

1. Learning activities demo
2. Search the Corpus
3. Create a Google site
4. Add embedding code for iDDL
5. Insert a collab platform URL
**Learning goals**

**Corpus input**

Interactive concordance lines

Target phrase highlighted

Sort by context before and after

**Interpretation questions**

Questions to guide pattern noticing

Inductive reasoning

**Practice activity**

Embedded Google Docs

Free production

Collaborative space
Example Activities demonstration
Go to sites.google.com and log in with your Google account

Click on Blank to create a new site

Give a title to your site

On the right hand side, locate the button Embed under the Insert tab

Click on the Pages tab to see the list of pages and subpages

While on the Pages tab, click on the plus button below to add new pages

Click on the three dots next to the Page name to create subpages
- Search for a word in Portuguese or Russian in MACAWS OR a word in English in CROW
- Use the filters on the left to select the language, course, genre, topics, draft, mode, and L1 background
- Click on Embed < /> to copy the code for your results that will be embedded in your activity
- Click on Copy to clipboard
- Go to your Google site
- Click on < > Embed on the left, select the Embed code tab, and paste the code
5

Insert a collab platform URL

- **Google Docs or Google Jamboard**
  - Go to docs.google.com
  - Click on **Blank** to create a new document
  - For public documents, click on **File>Publish to the Web**
  - Click on **Share** on the upper right
  - Click on **Get Link** to edit the privacy settings
  - Click on **Copy Link**
  - Go back to your Google Site and click on **Insert**, then `< > Embed`, then **By URL**
  - Paste the link and click on **Insert**

- **Padlet:**
  - Click on **Insert**, then `< > Embed`, then **Embed Code**
  - Paste the link and click on **Insert**
Questions?

Write your question in the “Q&A”
Your turn!

Practice the creation of your activities by searching these terms:

- “todos os” for intermediate Portuguese
- “руский язык” vs “по-русски” for beginner’s Russian
- “important” vs. “significant” for academic English writing

What questions would you ask to inductively guide students to discover patterns?
MACAWS blog with activities:  
https://sites.google.com/email.arizona.edu/macawswebinar/language-learning-activities

Crow blog with activities:  
https://writecrow.org/crow-for-teachers/

How to embed iDDL into LMS, sample materials, webinars, tutorials:  
http://macaws.flock.group

How to embed iDDL to Google Sites:  
https://youtu.be/Lo1Mz5LiElQ

iframe generator to embed Google Docs & others:  
https://www.iframe-generator.com/

Fill in the blanks activity generator:  

Articles

