Teaching STEM undergraduates discipline-specific writing skills: A data-driven learning approach

Ryan K. Boettger
University of North Texas

Stefanie Wulff
University of Florida

Introduction

Our research seeks to improve the quality of technical writing instruction for undergraduate STEM students. Specifically, we propose a data-driven learning (DDL) approach for teaching STEM students writing patterns specific to their respective disciplines.

DDL contrasts with traditional deductive, lecture-based methods and promotes students' active engagement through computer-assisted tools and a databank of authentic writing. Instruction is typically designed to –

- foster students' active engagement,
- encourage inductive activities that allow students to explore a topic on their own terms,
- promote interaction between students, and
- provide students with output-focused activities to apply this new knowledge [1].

Methods

We developed resources for six DDL units.

The first unit showed students how to perform basic functions in AntConc. AntConc is a free text processing tool that assists in the analysis of large data samples; it was specifically designed to facilitate DDL [2]. The five remaining units focused on a specific language patterns (see sidebar).

Testing Sites and Research Questions

The DDL treatment has been administered to over 200 participants with an additional 120 participants estimated for 2019. Each phase investigates a new question.

FALL 2018 — Taught in a 4000-level fish ecology course ($N = 17$)
- Does DDL work as a plug-in component in a course not focused on teaching writing?

SPRING 2019 — Trained 6 instructors to teach DDL in their technical writing courses ($N = 200$)
- Does DDL work with instructors new to the approach?

FALL 2019 — Co-teaching a CSE capstone course ($N = 120$ est.)
- Does DDL work (better) with more functionality and in larger class settings?

Data-Driven Learning Structure

Each treatment lasted 25-30 minutes and lessons followed the same structure. Students engaged in DDL activities in their class and then independently completed homework.

IN CLASS
- presentation
- 2 DDL activities (paper- and computer-based)

@HOME
- infographic fact sheet for reference
- 2 homework activities (one receptive, one productive skills)
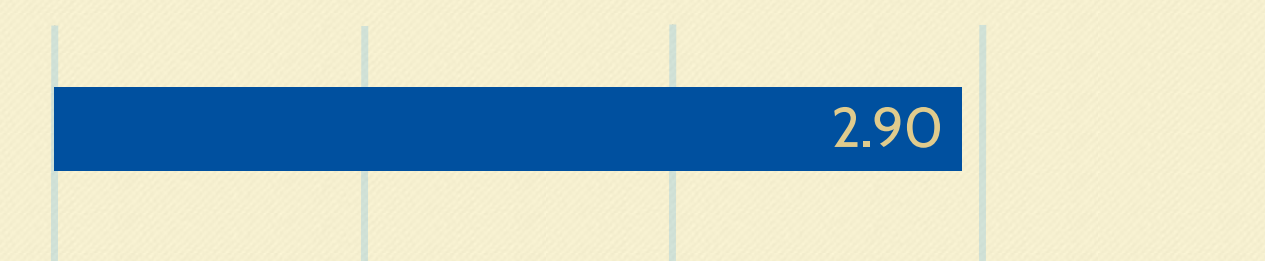
RESOURCES
- lesson plan
- activity and homework answer keys
- [lesson videos]
- [video DDL tutorials]

Results

Data from students' writing are currently being analyzed along various dimensions. One analysis will focus on students' uptake and change of use of the illustrative language examples presented in the instructional units and accompanying materials. This will be done by creating concordances and frequency counts of all attestations of the language examples in the student data (using the statistical and programming environment R).

In the final week of each semester, students completed a survey that collected their perceptions of the DDL units. Statements were ranked on a 4-point Likert scale (strongly disagree to strongly agree). Below are the current results ($N = 106$).

Statement	Score
I enjoyed the DDL in this course.	2.90
I understood the purpose of engaging in DDL in this course.	3.08
DDL taught me new approaches for improving my writing.	3.07
My understanding of language patterns benefited from the DDL.	2.90
I will use the data-driven-related skills I learned for my future writing.	2.98

Data-Driven Learning Units

Participants ranked the DDL units from most enjoyable to least enjoyable (or lowest to highest statistical mean).

1. Transitions ($M = 2.44$)

- **Therefore**, various life stages can be found in freshwater, brackish and marine habitats (Helfman et.al 2009, McDowall et. al 1981).

2. Hedges and Boosters ($M = 2.51$)

- …suggesting that lipids and carbohydrates **may** be more important for meeting their energy demands (Singer & Ballantyne, 2004).
- The river stingray is not a well-known animal to the public, but if its awareness is increased, it can **prove** to be a valuable tool to the future of freshwater conservation.

3. Evaluative Adjectives ($M = 3.23$)

- This may play an **important** role in regulating environmental impacts of invasive mollusks (Warren et al 2014).

4. Reporting Verbs ($M = 3.33$)

- There is also evidence to **suggest** that they use electrolocation through ampullary organs on their snout to sense prey (Jorgensen et al).

5. Passives ($M = 3.49$)

- The tone of the memo is very direct which could **be interpreted** as targeting and aggressive.

References

[1] K. Chujo, L. Anthony, and K. Oghigian, "DDL for the EFL classroom: Effective uses of a Japanese-English parallel corpus and the development of a learner-friendly, online parallel concordancer," in *American Association for Corpus Linguistics*, 2009.

[2] L. Anthony, "AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom," in *Professional Communication Conference Proceedings*, 2005, pp. 729–737.