

Crow: the Corpus & Repository of Writing

Components

- **Corpus of learner texts** from five writing genres (literacy autobiography, proposal, interview report, synthesis and research argument) totalling 6,727 texts of 6,394,882 words.
- **Repository of pedagogical artifacts** (syllabi, assignment sheets, assessment rubrics, activity sheets grammar worksheets, and group work activities).

Assignment Sequence	# of texts	# of words	Average word count
1. Literacy Autobio	1,249	1,474,747	1180.74
2. Proposal	1,638	1,311,472	800.65
3. Interview Report	676	590,058	872.87
4. Synthesis	1,554	1,458,339	938.44
5. Research Argt	1,610	1,560,266	969.11
TOTAL	6,727	6,394,882	950.63

Goals

- **Research:** provides data sets to examine student writing and classroom based instruction using corpus tools
- **Professional development:** a sustainable platform for writing instructors to share work with others
- **Mentoring:** provides mentors an archive of sample learner writing and assignments to share with their new TAs
- **Assessment:** supports implementation of common assignment outcomes across 100+ of FYW sections (L1 + L2 Writing)
- **Graduate-student driven:** promotes involvement of graduate students in research and project development
- **Sustainability:** from day one, embed sustainable practices in project decision-making, development, and research design.

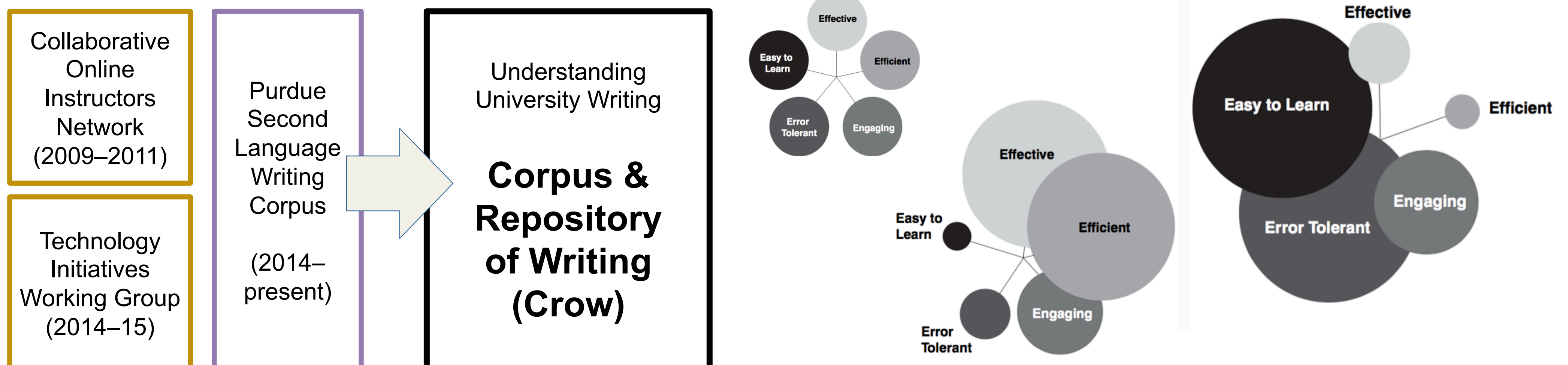
Team

- **Lead institutions:** Purdue University, University of Arizona and Michigan State University.
- **Partner institutions:** Northern Arizona University and University of South Carolina
- **Personnel:** Undergraduate students, graduate students, and faculty from applied linguistics, composition, and technical communication.

Wendy Jie Gao, Hadi Banat, Ge Lan, and Bradley Dilger, Purdue University

Development principles

- Flexible, user-centered approach (e.g. Quesenbery 2004). Acknowledge very wide ranges of users and tasks.
- Design for interdisciplinary, inter-institutional use.
- Researchers, developers, UX designers collaborate directly.



Plan of work

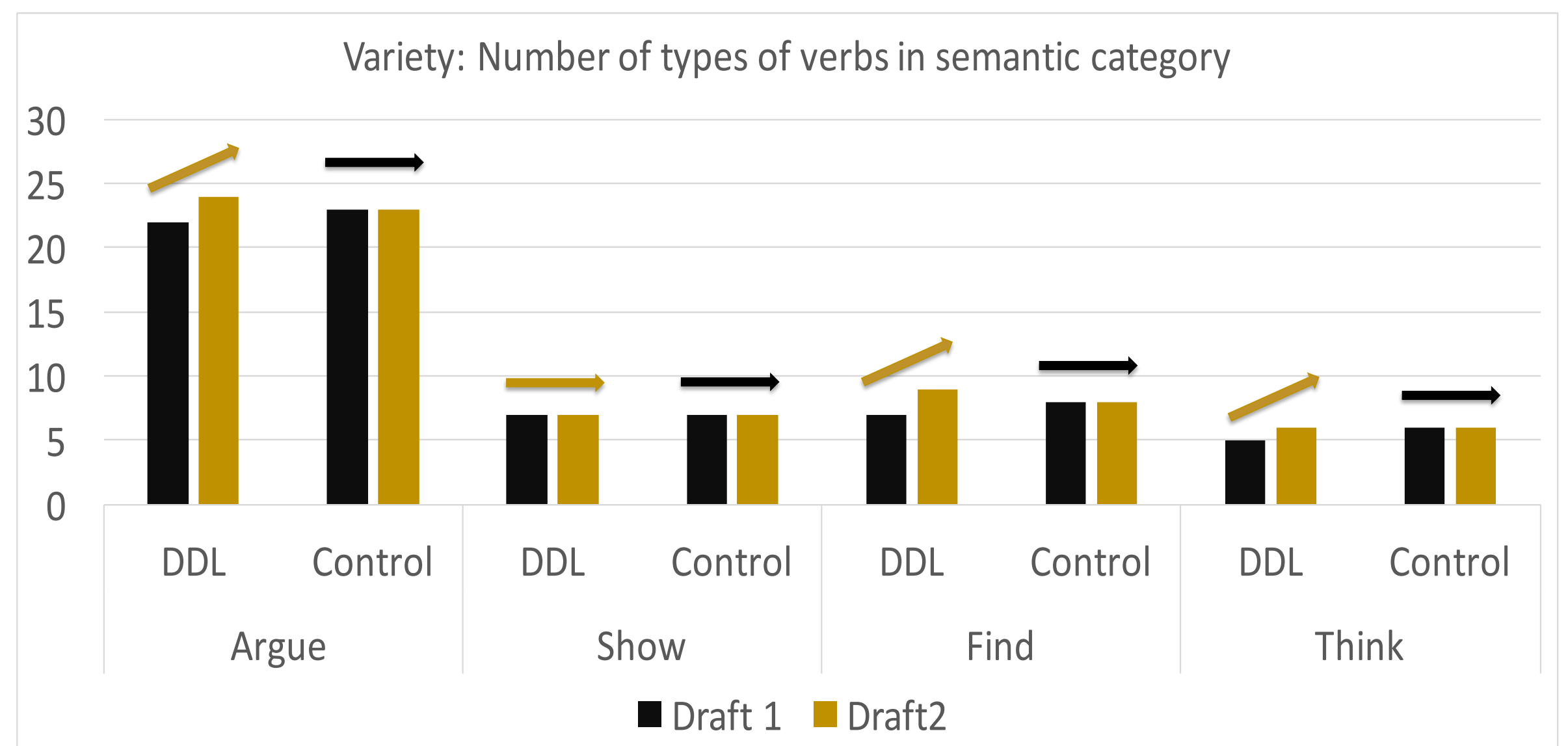
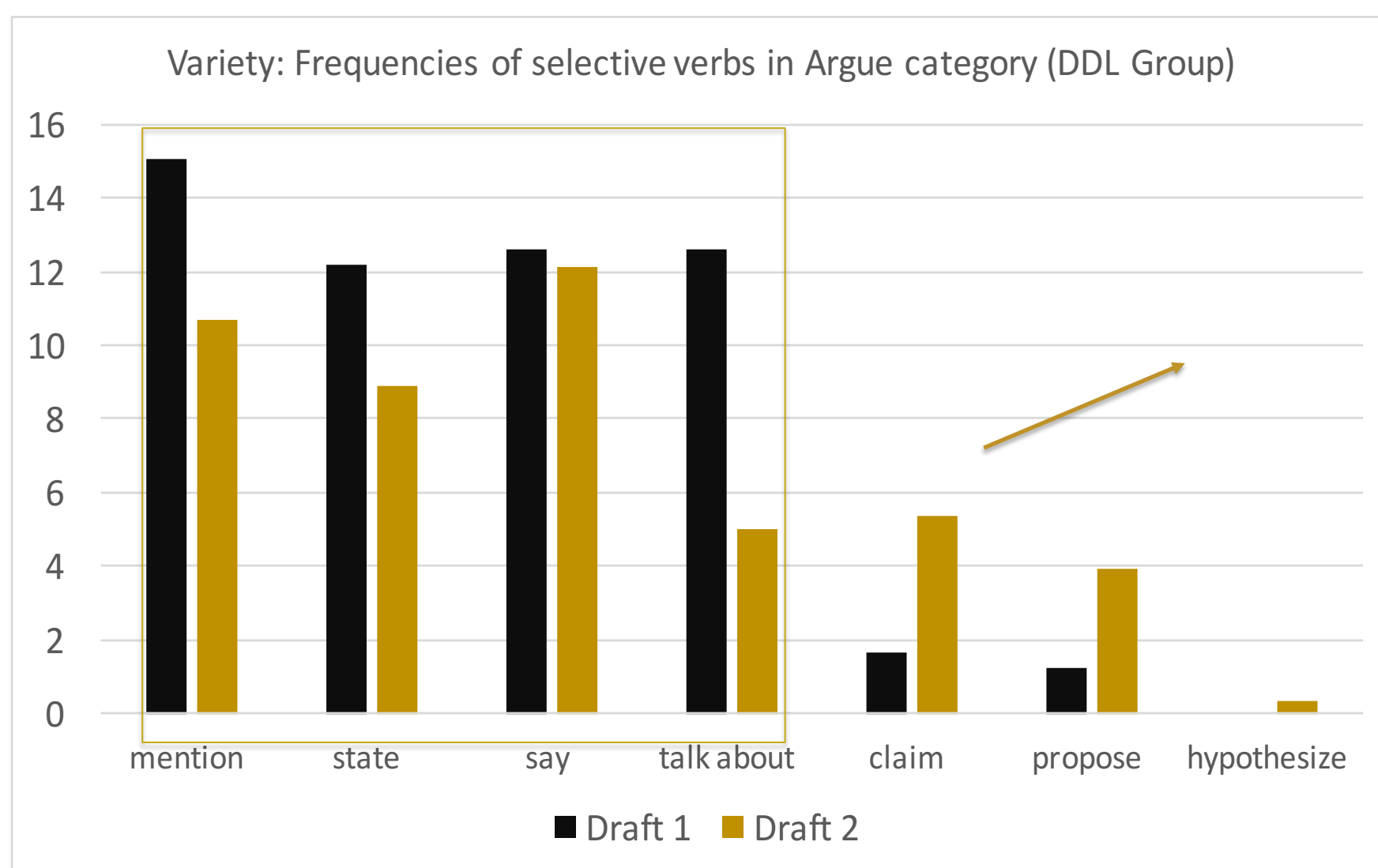
- Build corpora or incorporate corpora and repositories from multiple partners (started Fall 2014; ongoing)
 - *Purdue and Arizona*: learner corpora from L2-focused FYW
 - *Purdue*: repository from large composition program
 - *South Carolina*: corpus from FYW, focus on revision
 - *Northern Arizona*: corpus of L1 & L2 FYW
- Environmental scans of existing and defunct corpora & repositories (started Spring 2016; updated regularly)
 - Main audiences
 - Navigation
 - Ease of help
 - Styling and design
 - Overall use for Crow
 - Purposes
 - Intuitive use
 - Technologies used
 - **Sustainability**

Project timeline

- Develop database scheme, architecture suitable for both corpus & repository (started Spring 2017; working prototype Fall 2017)
- Develop simple web interfaces designed to support specific research projects (begin Summer 2017; working prototype Spring 2018)
- Alpha release corresponding with Fall 2018 symposium in West Lafayette (funded by HWW)

Reporting Verbs (Staples, Shin, Velásquez, Swatek, & Partridge, 2017)

- **Reporting verbs** provide attribution to sources (Charles, 2006; Francis, Hunston, & Manning, 1996; Swales, 2014).
- **Semantic Categories:** Argue, Show, Find, and Think.
- **Data Driven Learning (DDL):** using hands-on or paper-based, teacher-led activities to provide learners with an exposure to authentic language use (Boulton, 2009, 2010).
- **Research Question:** Will a DDL workshop in an L2 first-year composition class be effective in increasing reporting verb variety within the semantic categories?
- **Research Design:**
 - **DDL Group** (n=40, Spring 2016): Wrote **first draft** of a synthesis paper; Participated in 45-minute **DDL workshop**; Practiced **identifying reporting verbs** by using texts from PSLW; **Briefly revised papers, wrote second draft.**
 - **Control Group** (n=239, Fall 2014-Fall 2015): Wrote **the first draft** of a synthesis paper; **Received regular** instruction in class; Wrote **the second draft.**



Citation project (Gao, Wang, Lan, Staples, & Dilger, 2017)

- **Research Question:** In a First-Year Composition class for L2 students:
 - a) Will citation practices vary in number across different assignments?
 - b) Will the writers show a preference for particular citation styles?
- **Analytical Framework:** Integral citation and non-integral citation
- **Research Design:** 40 assignments were examined.
 - 20 are Literature Review and 20 are Research Report;
 - 20 come from the class of Instructor 1 and 20 come from Instructor 2.
 - Citation counts were highly variable, with 44% of writers using more than 10 citations and 33% using less than 3 per paper.

