



# Joel Dunham

---

## Education

- 2014 **PhD in Linguistics**, *The University of British Columbia*, Vancouver.  
Dissertation, entitled *The Online Linguistic Database: Software for Linguistic Fieldwork*, involved designing, coding, documenting, and arguing for innovative and open-source software to assist academics and other stakeholders in the collaborative documentation and analysis of under-studied and endangered natural languages.
- 2011 **Summer Institute, Linguistic Society of America**, *U. of Colorado*, Boulder.  
Won fellowship to take courses from international leaders in the fields of treebanking, statistical syntactic parsing, inference-based semantics, and computational lexical semantics.
- 2006 **Bachelor of Arts**, *The University of British Columbia*, Vancouver.  
Major in Philosophy, minor in Linguistics.

## Experience

- 2016–Present **Software Developer**,  
ARTEFACTUAL SYSTEMS INC., New Westminster.
- Designed, built, enhanced, and/or maintained web applications, web services, libraries, test and deployment infrastructure, and front-end components for *Archivematica*, a distributed application (and related tools) that performs preservation tasks on digital content and standards-compliantly documents the resulting metadata.
  - Recognized the need for better end-to-end testing and built a library of automated acceptance tests using Gherkin syntax, the Python `behave` library and Selenium to describe expected behaviour in non-technical but executable *feature files*; this significantly increased test coverage and was adopted by the DevOps team in their continuous integration testing and packaging pipelines.
  - Profiled our distributed task process-forking architecture using PyFlame, and as a result proposed and designed a forking reduction strategy which was a necessary cause in achieving a decrease in processing time of up to 80%.
  - Implemented a dependency-injection-based plugin API for METS Reader-Writer—an in-house library used to parse and create METS and PREMIS XML—, then composed and argued for the acceptance of an *Architectural Decision Record* which mandates the use of METS-RW in all future work; this simplified metadata serialization.
  - Communicated with international clients at major universities and libraries, refining requirements, generating estimates, performing code review, and planning and iterating over sprints.

#110, 1701 130<sup>th</sup> Street – Surrey, B.C. V4A 4A2

☎ (604) 838 4242 • ✉ [jrwdunham@gmail.com](mailto:jrwdunham@gmail.com)

🌐 [www.jrwdunham.com](http://www.jrwdunham.com) •

- 2014–Present **Full-stack Software Developer (Postdoctoral Fellow)**,  
CMLL DEPARTMENT, CONCORDIA UNIVERSITY, Montréal.
- Created SPAs, REST web services, administrative apps, migration scripts, NLP components, and deployment tools for *DativeBase*, an open-source software suite for documenting and analyzing endangered natural languages.
  - Designed, coded, and deployed *Dative*, a CoffeeScript/Backbone SPA that helps linguistic researchers to document and analyze under-studied languages. (GitHub repo)
  - Developed the *OLD*, Python (Pylons/Pyramid) software for building RESTful web services for managing collaborative linguistic databases. (Pylons repo, Pyramid repo)
  - Deployed 50+ language-specific *Dative/OLD* instances for academic research groups at institutions including UBC, McGill, the University of Helsinki, and UC Berkeley; wrote install and deploy tools using *Vagrant/Ansible* and *Docker (Compose)*, wrote documentation, provided technical support and performed custom data migrations.
  - Built abstractions over and integrations with NLP tools to help linguists create, configure and use finite-state phonologies and morphological parsers, and *N*-gram language models.
  - Managed contributions from other programmers via git and GitHub.
  - Presented, demoed and gave tutorials on the software internationally at major conferences in the field (LSA, ICLDC, etc.) and invited talks at Harvard, UC Berkeley and U Calgary.
  - Contributed to *FieldDB/LingSync*, open-source software for collaborative linguistic fieldwork (JavaScript, Node.js, AngularJS, Backbone, CouchDB.)
- 2011 & 2013 **Teaching Assistant**, THE UNIVERSITY OF BRITISH COLUMBIA, Vancouver.
- Assisted the teaching of *Linguistics 100: Introduction to Language and Linguistics*: led tutorials, held office hours, graded assignments and exams, and gave a lecture on computational fieldwork.
  - Assisted the teaching of *Linguistics 300: Studies in Grammar*: gave two lectures, held office hours, and graded assignments and exams.
- 2010–2012 **Linguistic Database Developer & Administrator**,  
LINGUISTICS DEPARTMENT, UNIVERSITY OF BRITISH COLUMBIA, Vancouver.
- Developed, deployed, administered, and provided technical support for several language-specific linguistic database web applications (*OLD v. 0.2*) used in field methods courses.
  - Applied for and won a SSHRC ITST grant to fund this innovative work.
  - Built the *OLD*, software for creating RESTful web service interfaces to linguistic databases (Python, Pylons, SQLAlchemy and MySQL.)
  - Wrote comprehensive unit and functional tests.
  - Incorporated open-source Natural Language Processing software for analytical and efficiency-promoting features:
    - *foma* for the construction of finite-state transducers for parsers,
    - *MITLM* for building morpheme sequence language models, and
    - *TGrep2* for performing structural searches over treebank corpora.
- 2005–2006 **Program Coordinator & Group Leader**,  
EXPLORE (SUMMER LANGUAGE BURSARY PROGRAM), Vancouver.
- Organized social and cultural activities for, and taught a writing class to, students learning French and English.

---

## Community Service

- Present **Chair of the Participation Committee**, COMMON GROUND CO-OP, Surrey.
- Document community participation credits and work to develop incentives for better participation.
  - As a board member and minutes secretary (2015–2017), helped in the conflict resolution, decision making, and management activities relevant to my housing co-operative.
  - Built a web application (Django, PostgreSQL) to track membership activities and archive meeting and committee minutes.
- Present **Scouter**, Scouts Canada, Surrey, BC.
- 2017 **Reviewer**, *Scientific Committee of The 11th International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan.
- 2016 **Reviewer**, *Program Committee of ComputEL-2, the 2nd Workshop on Computational Methods for Endangered Languages*, Honolulu, HI.
- 2008–2014 **Editor**, UBC WORKING PAPERS IN LINGUISTICS, Vancouver.
- Reputation for meticulous and constructive proofreading reflected by an acknowledgement in the text of one professor's paper (Matthewson 2011).

---

## Programming & Computer Skills

- Highly Productive PYTHON (DJANGO, PYRAMID, PYLONS, SQLALCHEMY), JAVASCRIPT (COFFEESCRIPT, BACKBONE, JQUERY), HTML, vim, git
- Productive CLOJURE, (CLOJURESCRIPT, RE-FRAME), SQL RDBMSs (MYSQL, POSTGRESQL, SQLITE), Linux/Unix, CSS,  $\LaTeX$ ( $X\LaTeX$ ,  $\text{Bib}\TeX$ ), FSTs (FOMA, XFST)
- Can Work With DOCKER (DOCKER COMPOSE), VAGRANT, ANSIBLE, BASH
- Basic Understanding ETHEREUM (SOLIDITY, WEB3.JS), COUCHDB, PERL, PYTHON TWISTED, PHP, REACTJS, ANGULARJS, GOLANG, JAVA

---

## Recent Presentations

- 2018 “On DativeBase, Archivemata and being a PhD outside of Academia”. Invited lecture for *Linguistics 518: Advanced Research Seminar*, UBC, Vancouver, CA, February 26.
- 2017 “Doing Right by their Data: a case study in digital preservation of domain-specific research data”. Panel presentation at the *Research Data Access & Preservation (RDAP) Summit*, Seattle, WA, April 20.
- 2016 “COULD Project Dative & the OLD”. Presentation at the *Digging Into Data Challenge Round Three Conference*, Glasgow, Scotland, January 28. (Co-presenter: Alan Bale)
- 2015 “LingSync: Software for language documentation”. Presentation at the *4<sup>th</sup> International Conference on Language Documentation & Conservation (ICLDC 4)*, Honolulu, Hawai'i, February 28.
- 2015 “LingSync Tutorial”. Presentation at the *Linguistic Society of America (LSA) Annual Meeting*, Portland, Oregon, January 8. (With Alan Bale and Jessica Coon.)

#110, 1701 130<sup>th</sup> Street – Surrey, B.C. V4A 4A2

☎ (604) 838 4242 • ✉ [jrwdunham@gmail.com](mailto:jrwdunham@gmail.com)

🌐 [www.jrwdunham.com](http://www.jrwdunham.com) •

- 2015 “LingSync: Software for collaborative linguistic fieldwork.” Presentation at the the Polinsky Language Sciences Lab meeting, Harvard Linguistics, May 6.
- 2015 “LingSync: Software for collaborative linguistic fieldwork”. Invited talk for the *Fieldwork Forum*, organized by the *Department of Linguistics at the University of California, Berkeley*, April 2.
- 2014 “The Online Linguistic Database: Software for Linguistic Fieldwork”. PhD dissertation oral defence, UBC, October 10.
- 2014 “LingSync & the OLD”. Presentation at the *Association for Computational Linguistics (ACL) Annual Meeting*, Baltimore, Maryland, June 18. (With Gina Chiodo.)

## Publications

- 2014 LingSync & the Online Linguistic Database: New models for the collection and management of data for language communities, linguists and language learners. In *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pp. 24-33, Baltimore, Maryland, USA. Association for Computational Linguistics. (Joel Dunham, Gina Cook, and Joshua Horner.)
- 2013 The Blackfoot Language Database. In K. S. Hele and J. R. Valentine (Eds.), *Papers of the Forty-First Algonquian Conference*, Concordia, Montreal, 2009, pp. 75-80.
- 2009 I qwlhí7censa sHenry: Henry’s shoes. In L. Matthewson, C. Christodoulou, J. Lyon & M. A. Oberg (eds.), *Wenácw Iz’: Sqwéqwel’s sLaura. True Stories by Laura Thevarge*, UBCWPL 22, pp. 31-35.
- 2007 The ‘durative’ in Blackfoot: understanding imperfectivity. In A.R. Deal (ed.), *SULA 4: Proceedings of the Fourth Conference on the Semantics of Under-represented Languages in the Americas*, Amherst: GLSA, pp. 49-64.
- 2006 Initial pitch informs sentence duration. In *The Journal of the Acoustical Society of America* 119, 3393. (Johanna Tan, Joel Dunham, Caleb Lee, and Eric Vatikiotis-Bateson.)

## Awards

- 2011 Linguistic Institute Fellowship, Linguistic Society of America
- 2011 Jacobs Research Funds Group Grant, Whatcom Museum Foundation (PI: J. Lyon)
- 2010 Image, Text, Sound & Technology Grant, Social Sciences & Humanities Research Council of Canada (PI: M. Wiltschko)
- 2010 Jacobs Research Funds Individual Grant, Whatcom Museum Foundation
- 2009 Jacobs Research Funds Group Grant, Whatcom Museum Foundation (co-applicant)
- 2007 Jacobs Research Funds Group Grant, Whatcom Museum Foundation (PI: M. Wiltschko)
- 2007 Canada Graduate Scholarship Doctoral Award, Social Sciences & Humanities Research Council of Canada
- 2006 Graduate Entrance Scholarship, University of British Columbia

#110, 1701 130<sup>th</sup> Street – Surrey, B.C. V4A 4A2

☎ (604) 838 4242 • ✉ jrwdunham@gmail.com

🌐 www.jrwdunham.com •

## Languages

English **First language**  
French **Intermediate**  
Latin **Basic**  
Blackfoot **Basic**

*Con conversationally fluent, able to read*  
*Basic grammatical knowledge and lexical recall*  
*Extensive fieldwork experience*

## Interests

- Piano & Guitar
- Cooking
- Mushroom hunting
- Running