

Blake H. Allen

CONTACT INFORMATION

21 Alvarado St., Apt. 7
Brisbane, CA 94005
USA

Phone: 415 756 8209
Email: allenblakeh@gmail.com
Website: <http://bhallen.com>
GitHub: <https://github.com/bhallen>

EMPLOYMENT

Yelp

Software Engineer, Search and Data Mining (June 2016 — present)

EDUCATION

University of British Columbia

Ph.D. from Department of Linguistics, November 2016

- Dissertation title: *Bayesian models of learning and generating inflectional morphology*
- Advisor: Prof. Douglas Pulleyblank

Harvard University

A.M. from Department of Linguistics, May 2011 (simultaneous A.M./A.B.)

- Thesis title: *A segmental basis for perceived syllabic and moraic structure*
- Advisor: Prof. Michael Becker

A.B. from Department of Linguistics (with Japanese as a Related Field), May 2011

- *summa cum laude*

SUMMARY OF EXPERTISE

- Proficient in **machine learning**, **statistical modeling**, and **natural language processing**
- 5 years of **software development** using **Python** and **JavaScript**, using **Git** for version control
- 4 years of **experimental design** and **data analysis in R**
- Solid computer science background, including **algorithms**, **data structures**, and **databases**
- Deep understanding of **formal linguistic analysis**
- 3 years of experience with **imaging data** and **speech signal processing**
- Highest-level certification in **Japanese**, working proficiency in **Mandarin**

MAJOR PROJECTS

ARIEL: Analysis of Rare Incident-Event Languages
(CMU Language Technologies Institute)

Feb. 2016 – May 2016

- *Goal*: a pipeline for low-resource language NLP development in response to emergent crises
- *My contributions*: data mining, creation of vectorized language representations for inference of properties of low-resource languages
- *Status*: currently in development as part of DARPA's Low Resource Languages for Emergent Incidents (LORELEI) project

Sublexical Morphology
(Dissertation research at UBC)

Jan. 2014 – November 2016

- *Goal*: a generalized probabilistic framework for learning and generating inflectional morphology
- *My contributions*: statistical model design, model implementation (Python), experimental design, data analysis, and interpretation (mixed effects models in R), corpus creation and munging
- *Status*: high accuracy on several datasets of up to approx. 1,000,000 wordforms

Phonological CorpusTools Apr. 2014 – June 2016
(UBC Department of Linguistics)

- *Goal*: an application that assists linguists by automating quantitative corpus analyses
- *My contributions*: analysis algorithm design and implementation (Python)
- *Status*: support for ~10 analysis types, compatibility with numerous corpus formats
- *Project website*: <http://phonologicalcorpustools.github.io/CorpusTools/>

Sublexical Learner for Morphophonology Jan. 2012 – Present
(UBC and Stony Brook University Departments of Linguistics)

- *Goal*: a learning algorithm for probabilistic models of phonology–morphology interactions
- *My contributions*: model design and implementation (JavaScript)
- *Status*: released with web app which includes dozens of example datasets
- *Project website*: <http://sublexical.phonologist.org/>

PhoMent: a Maximum Entropy phonotactic learner in Python Jan. 2014 – June 2016
(UCLA Department of Linguistics)

- *Goal*: a Python implementation of the UCLA learner for language-wide sound patterns
- *My contributions*: model design and implementation (Python)
- *Status*: currently in beta
- *Project website*: <https://github.com/rdaland/PhoMent>

eNunciate: online language learning technologies Apr. 2015 – June 2016
(UBC Centre for Teaching, Learning, and Technology)

- *Goal*: a web app that facilitates pedagogy of phonetics and second language pronunciation
- *My contributions*: biofeedback (speech signal processing) web app design, pedagogy site design, pedagogical materials design
- *Status*: currently in beta
- *Project website*: <http://enunciate.arts.ubc.ca/>

SUBMITTED
MANUSCRIPTS

(in review) **Allen, B.**, & Becker, M. Learning alternations from surface forms with sublexical phonology.

My contributions: algorithm design and implementation (JavaScript), dataset creation, writing

KEY: **: peer-reviewed by paper *: peer-reviewed by abstract

PUBLICATIONS

** Hall, K. Currie, **Allen, B.**, McMullin, K., Smith, H., & Yamane, N. (2017). Using optical flow analysis on ultrasound of the tongue to examine phonological relationships. *Canadian Acoustics*, 45(1), 15–24.

My contribution: ultrasound data analysis (Python and Smoothing Spline ANOVA in R)

** Gick, B., **Allen, B.**, Roewer-Després, F., & Stavness, I. (2017). Speaking tongues are actively braced. *Journal of Speech, Language, and Hearing Research*, 60, 494–506.

My contributions: electropalatography data analysis and visualization (R), writing

Allen, B., Archangeli, Diana, & Pulleyblank, D. (accepted). Assessing an allomorphy-based account of Japanese verb morphology. In Bouchard, D., Baronian, L., & Dubé, J. (Eds.), *The Simplicity of Grammar*. Oxford: Oxford University Press.

My contributions: algorithm design, corpus creation, all coding (Python), writing

** **Allen, B.**, Pulleyblank, D., & Ajíbóyè O. (2013). Articulatory mapping of Yoruba vowels: an ultrasound study. *Phonology*, 30(2), 183–210.

My contributions: data processing (Python) and analysis (Smoothing Spline ANOVA in R),

theory, writing

Allen, B. (2013). Hiatus-conditioned accent shift in Japanese: expanding the vowel sonority hierarchy. *UBC Working Papers in Linguistics*, 35.

Allen, B. (2013). Using computational models of grammar to assess lexical influences on experimental judgments. *UBC Working Papers in Linguistics*, 35.

** Pulleyblank, D., & **Allen, B.** (2013). Issues in the phonology-phonetics interface in African languages. In Orié, O., & Sanders, K. (Eds.), *Selected Proceedings of the 43rd Annual Conference on African Linguistics* (1–15). Somerville, MA: Cascadilla.

My contributions: literature review, publication statistics, writing

INVITED
CONFERENCE
PRESENTATIONS

* **Allen, B.** (2016). Sublexical Morphology: learning and generalizing probabilistic inflectional morphology. Workshop presentation. Workshop on Computational Methods for Descriptive and Theoretical Morphology, International Morphology Meeting. Vienna, Austria: Mid-February, 2016.

* **Allen, B.** & Hansson, G.Ó. (2016). Multiple bases and empirical priors in paradigm inference: experimental evidence from Icelandic. Paper presentation. International Morphology Meeting. Vienna, Austria: Mid-February, 2016.

My contributions: experimental design and analysis (R)

* **Allen, B.** (2015). Justified naivety: limits on constraint conjunction in inflectional morphology. Poster presentation. Annual Meeting on Phonology. Vancouver, BC: October 10, 2015.

* Yamane, N., Abel, J., **Allen, B.**, Burton, S., Kazama, M., Kim, B., Noguchi, M., Tsuda, A., & Gick, B. (2015). Ultrasound-enhanced multimodal approaches to pronunciation teaching and learning. Paper presentation. Acoustics Week in Canada. Halifax, NS: October 7, 2015.

My contributions: biofeedback web app design, pedagogy site design, pedagogical materials design

* Currie Hall, K., **Allen, B.**, Fry, M., Mackie, S., & McAuliffe, M. (2015). Calculating functional load with pronunciation variants. Workshop presentation. Workshop on Modeling Variability in Speech, Institute for Natural Language Processing, University of Stuttgart. Stuttgart, Germany: October 2, 2015.

My contributions: algorithm design and implementation (Python)

* **Allen, B.** (2015). Bayesian morphology: simplifying the challenge of productivity. Poster presentation. Workshop on Computational Phonology and Morphology, LSA Summer Institute. Chicago, IL: July 11, 2015.

* McMullin, K., & **Allen, B.** (2015). Phonotactic learning and the conjunction of Tier-based Strictly Local languages. Paper presentation. 89th Meeting of the Linguistic Society of America. Portland, OR: January 10, 2015.

My contributions: algorithm design, all coding (Python)

* **Allen, B.**, & Becker, M. (2015). Learning alternations from surface forms with sublexical phonology. Poster presentation. 89th Meeting of the Linguistic Society of America. Portland, OR: January 9, 2015.

My contributions: algorithm design, most coding (JavaScript)

* **Allen, B.** (2014). Evaluating distance-based alternatives to natural class-based constraints. Poster presentation. Annual Meeting on Phonology 2014. Cambridge, MA: September 21, 2014.

- * **Allen, B.**, & Becker, M. (2014). Learning alternations from surface forms with sublexical phonology. Paper presentation. Annual Meeting on Phonology. Cambridge, MA: September 19, 2014.
My contributions: algorithm design, most coding (JavaScript)
- * Currie Hall, K., Smith, H., McMullin, K., Yamane, N., **Allen, B.**, & Gambarage, J. (2014). Articulatory correlates of phonological relationships. Paper presentation. 14th Conference on Laboratory Phonology. Tokyo, Japan: July 27, 2014.
My contributions: ultrasound consultation and data organization, experiment design
- * Mackie, Scott, Currie Hall, K., **Allen, B.**, Fry, M., McAuliffe, M., and McMullin, K. (2014). Phonological CorpusTools: a free, open-source tool for phonological analysis. Poster presentation. 14th Conference on Laboratory Phonology. Tokyo, Japan: July 26, 2014.
My contributions: algorithm design, command line interface design (Python)
- * Gick, B., **Allen, B.**, Stavness, I., & Wilson, I. (2013). Speaking tongues are always braced. Poster presentation. 166th Meeting of the Acoustical Society of America. San Francisco, CA.
My contributions: EPG data aggregation and analysis (R)
- * **Allen, B.** (2013). A computational model of grammar inference for assessing lexical influences on experimental judgments. Poster presentation. Annual Conference of the Canadian Linguistic Society. Victoria, BC.
- * **Allen, B.** (2013). Lyman’s Law and limits on “leaky” phonotactics. Paper presentation. The 29th Northwest Linguistics Conference. Vancouver, BC.
- * **Allen, B.** & Noguchi, Masaki (2013). Japanese accent shift: generalizing from ambiguous input. Paper presentation. 87th Meeting of the Linguistic Society of America. Boston, MA.
My contributions: experiment design, data collection, data analysis, theory
- * Pulleyblank, D., & **Allen, B.** (2012). Issues in the phonology-phonetics interface in African languages. Plenary talk. 43rd Annual Conference on African Linguistics. New Orleans, LA. Invited oral presentation by D. Pulleyblank.
My contributions: literature review, publication statistics

HONORS AND
AWARDS

- St. John’s College Sir Quo-Wei Lee Fellowship Fall 2015
◦ support for research-related travel and thesis completion
- UBC Teaching and Learning Enhancement Fund Research Assistantship Summer 2015 – June 2016
- Walter H Gage and Elsie M Harvey Education Abroad Scholarship Winter 2013/4
◦ support for study abroad at UCLA with Prof. Bruce Hayes
- Research Assistantship with K. Currie Hall (UBC) Winter 2013 – June 2016
- Linguistic Society of America Summer Institute Fellowship Summer 2013
◦ support for attending the LSA 2013 Summer Institute
- Research Assistantship with D. Pulleyblank (UBC) Winter 2011/2 – Summer 2015
- Research Assistantship with Gunnar Ó. Hansson (UBC) 2011 – 2012, 2015 – June 2016
- Four Year Fellowship (UBC) Fall 2011 – Summer 2015

	International Partial Tuition Scholarship (UBC)	Fall 2011 – Summer 2015
	Membership, Phi Beta Kappa Society (Harvard University Alpha Iota)	Spring 2011 – Present
	◦ inducted into Harvard University’s chapter, Alpha Iota of Massachusetts	
	Harvard College Scholar Award	Spring 2010
	◦ awarded to students whose GPA is in the highest 5% to 10%	
	Reischauer Institute Undergraduate Summer Language Study Grant	Spring 2009
	◦ support for participation in Princeton in Ishikawa (see below)	
	David Rockefeller International Experience Grant	Spring 2009
	◦ support for participation in Princeton in Ishikawa (see below)	
	Harvard College Research Program Grant	Spring 2008
	◦ support for field research with Prof. Andrew Nevins	
OTHER	Visiting student at UCLA	Jan. – May 2014
EDUCATION	◦ regular research meetings with Prof. Bruce Hayes; attended Computational Phonology course taught by Prof. Robert Daland	
	Linguistic Society of America 2013 Summer Institute	Summer 2013
	◦ courses: Machine Learning, Quantitative and Computational Phonology, Mixed Effects Models, Computational Modeling of Sound Change, Phonological Acquisition	
	Harvard Beijing Academy: Intermediate Mandarin	Summer 2011
	◦ 9-week program in Beijing: studied Mandarin grammar and Chinese culture equivalent to two semesters at the 2nd-year level	
	Kyoto Consortium for Japanese Studies: Classical Japanese	Summer 2010
	◦ 6-week program at Dôshisha University in Kyoto: studied introductory Classical Japanese grammar and literature equivalent to two semesters	
	Princeton in Ishikawa: Advanced Modern Japanese	Summer 2009
	◦ 8-week program in Kanazawa, Ishikawa prefecture: studied Modern Japanese grammar and literature equivalent to two semesters at the 3rd-year level	
NATURAL	◦ English (native speaker)	
LANGUAGE	◦ Japanese (fluent: JLPT N1 certificate)	
FLUENCIES	◦ Mandarin (intermediate: four semesters with study abroad)	
	◦ Latin (reading knowledge: 5 years)	
	◦ Classical Japanese (reading knowledge: 2 semesters)	