

Curriculum Vitae

December 7, 2015

David R. Mortensen
Language Technologies Institute
Carnegie Mellon University
Pittsburgh, PA 15213
United States of America

Academic Employment

- 2015–2018 Research Scientist, Language Technologies Institute, Carnegie Mellon University
- 2014–2014 Visiting Researcher, Language Technologies Institute, Carnegie Mellon University
- 2006–2014 Assistant Professor, Department of Linguistics, University of Pittsburgh
- 2003–2004 Graduate Student Instructor, Department of Linguistics, University of California, Berkeley.

Educational Background

Education

- B.A. in English Literature (Departmental Honors; minors in Linguistics and Psychology). 2000. Utah State University. Honors Thesis: *Sinitic Loanwords in Two Hmong Dialects of Southeast Asia*.
- M.A. in Linguistics. 2003. University of California, Berkeley.
- LSA Summer Institute. 2006. Massachusetts Institute of Technology.
- Ph.D. in Linguistics. 2006. University of California, Berkeley. Dissertation: *Logical and Substantive Scales in Phonology*, supervised by Sharon Inkelas and James A. Matisoff (co-chairs), Larry Hyman and Johanna Nichols (committee members).

Languages

Mong Leng Advanced (fluent) reading, writing, and speaking ability.
Hmong Daw Advanced reading knowledge and intermediate speaking ability.
Mandarin Intermediate reading knowledge.
French Elementary reading ability.
German Elementary speaking and reading ability.

Technical Skills

Programming languages Python, Perl, Ruby, R, Javascript, Haskell, Java

Markup languages XML, LaTeX
Other technologies SQL, Unicode

Research Interests

- Linguistic typology.
- Phonology and its relationship to phonetics and morphology.
- Morphology and the lexicon.
- Computational methods in theoretical, descriptive, and historical linguistics.
- Annotation and data resource creation.
- Southeast Asian linguistics (Hmongic, Tibeto-Burman).
- Language change and historical linguistics.

Fellowships, Scholarships, and Academic Awards

- (2010) University of Pittsburgh Arts and Sciences Faculty Grants Committee Summer Research Grant, Type I.
- (2006) LSA Summer Institute Fellowship
- (2004) Graduate Division Summer Grant, University of California, Berkeley
- (2001) Institute for East Asian Studies Fellowship, University of California, Berkeley

Publications and Other Research Products

Refereed Journal Articles

- Littell, Patrick, Tian Tian, Ruochen Xu, Zaid Sheikh, David Mortensen, Lori Levin, Francis Tyers, Hiroaki Hayashi, Graham Horwood, Steve Sloto, Emily Tagtow, Alan Black, Yiming Yang, Teruko Mitamura, and Eduard Hovy (2017). “The ARIEL-CMU situation frame detection pipeline for LoReHLT16: a model translation approach.” *Machine Translation*: 1–22. DOI: 10.1007/s10590-017-9205-3
- Mortensen, David R. (2013). “Tonally Conditioned Vowel Raising in Shuijingping Mang.” *Journal of East Asian Linguistics* 22(3): 189–216. DOI: 10.1007/s10831-013-9102-6
- Mortensen, David R. and James A. Miller (2013). “A Reconstruction of Proto-Tangkhuic Rhymes.” *Linguistics of the Tibeto-Burman Area* 36(1): 1–32.
- Mortensen, David R. (2012). “The Emergence of Dorsal Stops after High Vowels.” *Diachronica* 29(4): 434–470. DOI: 10.1075/dia.29.4.02mor
- Vercellotti, Mary Lou and Mortensen, David R. (2012). “A Classification of Compounding in American Sign Language: an Evaluation of the Bisetto and Scalise Framework.” *Morphology* 22(4): 545-579. DOI: 10.1007/s11525-012-9205-1

Mortensen, David R. and Keogh, Jennifer (2011). “Sorbung, an Undescribed Language of Manipur: Its Phonology and Place in Tibeto-Burman” *Journal of the Southeast Asian Linguistics Society* 4(1): 62–114.

Refereed Conference Proceedings

- Littell, Patrick; McCoy, Tom; Han, Na-Rae; Rijhwani, Shruti; Sheik, Zaid; Mortensen, David; Mitamura, Teruko and Levin, Lori (forthcoming) “Parser combinators for Tigrinya and Oromo morphology,” *Proceedings of the Twelfth International Conference on Language Resources and Evaluation* (LREC 2018).
- Mortensen, David R.; Dalmia, Siddharth; and Littell, Patrick (forthcoming) “Epitran: precision G2P for many languages,” *Proceedings of the Twelfth International Conference on Language Resources and Evaluation* (LREC 2018).
- Littell, Patrick; Mortensen, David; Lin, Ke; Kairis, Katherine; Turner, Carlisle and Levin, Lori (2017) “URIEL and lang2vec: Representing languages as typological, geographical, and phylogenetic vectors,” *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 8–14, Valencia, Spain, April 3–7, 2017 (EACL 2017).
- Littell, Patrick; Goyal, Kartik; Mortensen, David; Little, Alexa; Dyer, Chris and Levin, Lori (2016) “Named Entity Recognition for Linguistic Rapid Response in Low-Resource Languages: Sorani Kurdish and Tajik”, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical*, 998–1006, Osaka, Japan, December 11–17 2016 (COLING 2016).
- Mortensen, David R.; Littell, Patrick; Bharadwaj, Akash; Goyal, Kartik; Dyer, Chris and Levin, Lori (2016) “PanPhon: A Resource for Mapping IPA Segments to Articulatory Feature Vectors”, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 3475–3484, Osaka, Japan, December 11–17 2016 (COLING 2016).
- Bharadwaj, Akash; Mortensen, David; Dyer, Chris and Carbonelle, Jaime (2016) “Phonologically aware neural model for named entity recognition in low resource transfer settings,” in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 1462–1472, Austin, Texas, November 1–5, 2016 (EMNLP 2016) .
- Littell, Patrick; Mortensen, David R.; Goyal, Kartik; Dyer, Chris and Levin, Lori (2016) “Bridge-language capitalization inference in Western Iranian: Sorani, Kurmanji, Zazaki, and Tajik”, in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (LREC 2016).
- Tsvetkov, Julia; Sitaram, Sunayana; Lample, Guillaume; Littell, Patrick; Mortensen, David; Black, Alan W.; Levin, Lori and Dyer, Chris (2016) “Polyglot neural language models: a case study in cross-lingual phonetic representation learning,” in *Proceedings of NAACL-HLT 2016*, 1357–1366, San Diego, California, June 12–17, 2016.

Other Academic Publications

- Mortensen, David R. (forthcoming). “Hmong-Mien Languages.” In Oxford Research Encyclopedia of Linguistics.
- Mortensen, David R. (forthcoming). “Hmong (Mong Leng).” Invited chapter to appear in *The Mainland Southeast Asian Linguistic Area*, to be published by de Gruyter Mouton (edited by Alice Vittrant and Justin Watkins).
- Mortensen, David R. (2017). “Hmong-Mien Languages.” In Oxford Bibliographies in Linguistics. Ed. Mark Aronoff. New York: Oxford University Press.
- Mortensen, David R. (2013). “Lexical Prefixes and Tibeto-Burman Laryngeal Contrasts.” *Proceedings of the Thirty-Seventh Annual Meeting of the Berkeley Linguistics Society* (BLS 37), 272–286.
- Mortensen, David R. (2004). “The emergence of dorsal stops after high vowels in Huihu.” *Proceedings of the Thirtieth Annual Meeting of the Berkeley Linguistics Society* (BLS 30), 292–303.
- Mortensen, David R. (2003). “Review of *Baheng Yu Yanjiu* [research on the Pa-Hng language] by Mao Zongwu and Li Yunbing.” *Linguistics of the Tibeto-Burman Area* 26(1): 123–131.
- Mortensen, David R. (2002). “Review of *Les langue Hmong-Mjen: phonologie historique* by Barbara Niederer,” *Linguistics of the Tibeto-Burman Area* 25(2): 243–245.

Data Resources

- Mortensen, David R. (2014). “SCH Corpus.” 15 million word unannotated Hmong-language corpus derived from postings to the soc.culture.hmong Usenet group. http://www.davidmortensen.org/corpora/sch_corpus-2.zip

Published Puzzles from Linguistic Olympiads

- Mortensen, David R. and Littell, Patrick (2013). “Rewrite me Badd.” In *Puzzles in Logic, Language, and Computation: The Red Book*, edited by Dragomir Radev, 9. Berlin: Springer.
- Mortensen, David R. (2013). “Hypo-Hmong-driac.” In *Puzzles in Logic, Language, and Computation: The Red Book*, edited by Dragomir Radev, 47. Berlin: Springer.
- Mortensen, David R. (2013). “Tangkhul Tangle.” In *Puzzles in Logic, Language, and Computation: The Green Book*, edited by Dragomir Radev, 13. Berlin: Springer.
- Mortensen, David R. and Patrick Littell (2013). “Tale of Kieu.” In *Puzzles in Logic, Language, and Computation: The Green Book*, edited by Dragomir Radev, 28. Berlin: Springer.

Unpublished Software

- Mortensen, David R. (2012). *NetSPE*. A web-based application for exploring rule based analyses of phonological data. <https://github.com/dmort27/NetSPE>
- Mortensen, David R. (2011). *Web Comparator*. A web-based application for organizing and analyzing comparative lexical databases. Used in the production of

“Emergence of Obstruents” and “Proto-Tangkhulic Rhymes” above.
<https://github.com/dmort27/WebComparator>

Mortensen, David R. (2006). *HangGram*. A tool for learning to read spectrograms of human speech through the use of a “hangman” paradigm. Used as a supplement to classroom teaching.

Presentations

Invited Talks

- Mortensen, David R. (2010). “Does Hmong allow noun incorporation? A field study,” colloquium talk, University of Pittsburgh, November 19.
- Mortensen, David R. (2010). “Does Hmong allow Noun Incorporation?” invited talk, Wayne State University, Detroit, October 1.
- Mortensen, David R. (2010). “Shuijingping Hmong and the problem of tone-vowel interactions,” invited talk, the University at Buffalo, February 12.
- Mortensen, David R. (2009). “The Problem of Tone-Vowel Interactions,” invited talk, Chulalongkorn University, Bangkok, November 10.
- Mortensen, David R. (2008). “Optimal-Constructional Morphology, or how, when our lexicon gives us lemons, we make lemonade,” colloquium talk, University of Pittsburgh, November 25.
- Mortensen, David R. (2006). “A Marked Structure is Born: The Emergence of Obstruents and the Motivations for Sound Change,” invited talk, University of Pennsylvania, Philadelphia, February 16.
- Mortensen, David R. (2005). “The Phonology of Coordinate Constructions,” invited talk, Cornell University, Ithaca, March 3.
- Mortensen, David R. (2005). “Abstract Scales in Phonology,” invited talk, Stanford Phonology Workshop, Stanford February 24.
- Mortensen, David R. (2005). “Abstract Scales in Phonology,” invited talk, University of Chicago, Chicago, February 3.
- Mortensen, David R. (2004). “Scalar Competition and Variables in Hmong Anaphora,” invited talk, Stanford Syntax Workshop, Stanford, April 27, 2004.
- Mortensen, David R. (2003). “The State of Hmong Linguistics,” invited talk, Hmong Studies Symposium, Concordia University (St. Paul), October 25, 2003.

Refereed Conference Presentations

- Mortensen, David R. (2014). “Learning phonological ordering generalizations for Hmong elaborate expressions,” Linguistic Society of America, Minneapolis, January 3.
- Mortensen, David R. (2013). “Tone and the order of Hmong elaborate expressions—a corpus-based investigation,” International Conference on Sino-Tibetan Languages and Linguistics 46, Hanover, New Hampshire, August 8.
- Mortensen, David R. (2011). “Two series suffice: Lexical prefixes and Proto-Tibeto-Burman laryngeal contrasts,” Berkeley Linguistics Society 37, Berkeley, California, February 12–13.

- Mortensen, David R. (2011). "Pseudo-Incorporation in Hmong," Linguistic Society of America, Pittsburgh, January 8.
- Mortensen, David R. (2010). "The Place of Sorbung in Tibeto-Burman," International Conference on Sino-Tibetan Languages and Linguistics 43, Lund, October 16.
- Mortensen, David R. and Keogh, Jennifer (2010). "A phonological sketch of Sorbung," Southeast Asian Linguistics Society 20, Zürich, June 11.
- Mortensen, David R. (2010). "Morphological doubling: Evidence from Jingpho echo-reduplication and coordinate compounding," Linguistic Society of America, Baltimore, January 10.
- Mortensen, David R. and James A. Miller (2009). "Proto-Tangkhul Onsets in Comparative Perspective," International Conference on Sino-Tibetan Languages and Linguistics 42, Chiangmai, November 4.
- Mortensen, David R. and James A. Miller (2009). "Proto-Tankhulic Rhymes," Southeast Asian Linguistics Society 19, Ho Chi Minh City, May 28-29.
- Mortensen, David R. (2006). "Coordinate Compound Ordering Effects and Abstract Phonological Scales," West Coast Conference on Formal Linguistics 25, Seattle, WA, April 28-30.
- Mortensen, David R. (2006). "Tonally conditioned vowel raising in Shuijingping Hmong," Linguistic Society of America, Albuquerque, New Mexico, January 5-8.
- Mortensen, David R. (2005). "Phonological Parallelism between Reduplication and Compounding in Jingpho," Berkeley Linguistics Society 31, Berkeley, California, February 18-20.
- Mortensen, David R. (2005). "A-Hmao echo reduplication as evidence for abstract phonological scales," Linguistic Society of America, Oakland, California. January 6-9.
- Mortensen, David R. (2004). "Phonological Cline Effects" (poster), Northeast Linguistics Society 35, Storrs, Connecticut, October 22.
- Mortensen, David R. (2004). "The emergence of dorsal stops after high vowels in Huihu," Berkeley Linguistics Society 30, Berkeley, California, February 14.
- Mortensen, David R. (2004). "*Semper Infidelis*: Anti-identity in A-Hmao and Jingpho tone sandhi chains," Linguistic Society of America 78. Boston, Massachusetts, January 10.
- Mortensen, David R. (2003). "Two types of variable elements in Hmong anaphora," Workshop on the Form and Function of Pronouns, University of British Columbia, September 19-21.
- Mortensen, David R. and Kenneth VanBik (2003). "Vowel quality and quantity in Hakha Lai: A perceptual study," Linguistic Society of America 77. Atlanta, Georgia, January 5.
- Mortensen, David R. (2002). "A preliminary survey of Tibeto-Burman loanwords in Hmong-Mien languages," International Conference on Sino-Tibetan Languages and Linguistics 35. Arizona State University.
- Mortensen, David R. (2002). "Borrowed animal terms in Hmong-Mien languages," Southeast Asian Linguistics Society 12. Northern Illinois University.

Other Conference and Workshop Presentations

Mortensen, David R. (2013). "Kuki-Chin Phonology: an Overview," International Conference on Sino-Tibetan Languages and Linguistics 46, Workshop on Issues in Kuki-Chin Linguistics, October 7.

Mortensen, David R. (2003). "Chain-Shift, Schmain-Shift: Anti-Identity and Tone Sandhi in Hmong, A-Hmao, and Jingpho," Trilateral Phonology Weekend 2003, University of California, Santa Cruz, May 10.

Teaching

Courses Taught

Carnegie Mellon University

11-411/11-611: Natural Language Processing S 2017, F 2017, S 2018

11-421/11-721: Grammars and Lexicons F 2017

University of Pittsburgh

(Courses 2000 and above are at the graduate level)

Ling 1441/2441: Field Methods S 2007, 2008, 2010, 2011, F 2012, S 2012, S 2014

Ling 1578: Phonetics and Phonemics F 2008, 2010, 2011

Ling 1579: Phonology S 2007, 2010, 2013

Ling 1773: Morphology F 2006

Ling 1951: Languages of the World F 2007, 2008, 2011, 2013

Ling 2578: Phonetics and Phonemics F 2006

Ling 2579: Phonology S 2008, 2010, 2012, 2013, 2014

Ling 2580: Topics in Phonological Theory S 2008

Ling 2773: Morphology F 2007, 2012, 2013

Ling 3773: Advanced Morphology S 2011

University of California, Berkeley

Outstanding Graduate Student Instructor Award 2004-2005

Ling 130: Historical and Comparative Linguistics
(TA)

Ling 5: Language and Linguistics (TA) F 2003

Supervising

PhD Students

Wang Ping (2015). Phonological prominence and its interaction with tone in Chinese dialects. PhD dissertation, University of Pittsburgh. (Co-chair: Shelome Gooden)

PhD Dissertation Committees

Marcos, Nausica (in progress). *Instruction and Learning of Derivational Morphology by English and German-speaking Learners of Spanish as a Foreign Language*. PhD dissertation. (Chair: Alan Juffs).

Friedline, Benjamin (2011). *Morphological Challenges in the Second Language Acquisition of Derivational Morphology*. PhD dissertation. (Chair: Alan Juffs).

Rodriguez, Guillermo (2008). *Second language sentence processing: Is it fundamentally different?* PhD dissertation. (Chair: Alan Juffs)

Eberhardt, Maeve (2009). *Identities and Local Speech in Pittsburgh: A Study of Regional African American English*. (Chair: Scott Kiesling).

MA Thesis Committees

Fan, Ya-wen (2013). *Gender Influences on the Use of Mandarin Varieties in Taiwan*. MA Thesis. (Chair: Scott Kiesling).

Cui Jie (2010). *Variation in Nanchang Gan*. MA Thesis. (Chair).

Draucker, Fawn (2010). *Revisiting the Maori English Vowel Space: Exploring variation in /I/ and /u/ vowel production in Auckland, New Zealand*. MA thesis. (Chair).

Friedline, Benjamin (2008). *A Linguistic Profile of Power and Identity in Massively Multiplayer Online Role-Playing Games*. (Chair: Scott Kiesling)

Collister, Lauren B. (2008). *Virtual discourse structure: an analysis of conversation in World of Warcraft*. MA thesis. (Chair: Scott Kiesling)

Littell, Patrick (2008). "Body part incorporation." MA long paper. (Co-chair with Pascual Masullo).

McCarthy, Natalie (2007). *The Northern Cities Shift and local identity in a suburban Cleveland group*. MA thesis. (Chair: Scott Kiesling).

Comprehensive Paper Committees

Tse, Holman (2013). "The emergence of retroflexion in Somali Bantu Kizigua: Internal motivation or contact-induced change?" (Chair)

Wang, Ping (2012). "The duration of medial glides in first and second language speakers of Mandarin."

Olsen, Michael (2012). "The Effect of Animacy and Pronominal Object Clitic Distinction in L2 Spanish." (Chair: Alan Juffs).

Marcos, Nausica. "An Analysis of Spanish Verb-Noun Compounds as Input for Derivation." (Chair: Alan Juffs).

Wang, Ping (2012). "An OT Account of Major Patterns of Chinese Coda Evolution." (Chair).

- Vercellotti, Mary Lou (2011). "A classification of compounding in American Sign Language and an evaluation of the Bisetto and Scalise Framework." Comprehensive paper. (Co-chair with Claude Mauk).
- Luk, Pei-Su (2011). "Investigating the transitive and intransitive constructions in English, Chinese, and Japanese from a functional perspective." (Chair: Yasuhiro Shirai).
- Collister, Lauren B. (2010). "Meaning Variation of the Iconic Deictics ^ and <- in an Online Community." (Co-chair with Scott Kiesling).
- Lifrieri, Vernica (2009). "A puzzle for effort in phonetically-grounded OT: codas in North-Central Peninsular Spanish." Comprehensive paper. (Co-chair with Shelome Gooden).
- Chan, Derek Ho-Leung (2009). "The tonal morphology of Cantonese attenuative reduplication." Comprehensive paper. (Co-chair with Shelome Gooden).
- Martelle, Wendy (2007). "The evolution of aspect: a case study in Russian." Comprehensive paper. (Chair: Yasuhiro Shirai)
- Pharmakis, Patricia (2007). "A formal account of Modern Greek inflection within the framework of Distributed Morphology." Comprehensive paper. (Chair: Alan Juffs)
- Brubaker, Brian (2007). "Implications of labial-tone correspondence in L2 Mandarin Chinese." Comprehensive paper. (Co-chair with Shelome Gooden).

Service

Editorial Boards

- *Journal of the Southeast Asian Linguistics Society*

Ad Hoc reviewer

Journals

- *Computational Linguistics*
- *Himalayan Linguistics*
- *Diachronica*
- *Language and Intercultural Communication*
- *Lingua*
- *Linguistics of the Tibeto-Burman Area*
- *Natural Language and Linguistic Theory*
- *Phonology*
- *Journal of the International Phonetic Association*
- *Journal of the Southeast Asian Linguistics Society*

Conferences

- Language Resources and Evaluation Conference (LREC)
- COLING

Granting agencies

- National Science Foundation

Departmental service

- Undergraduate advisor, Department of Linguistics, University of Pittsburgh. (Spring 2008–present).
- Colloquium committee, Department of Linguistics, University of Pittsburgh (Fall 2007–2009).
- Curriculum committee, Department of Linguistics, University of Pittsburgh (2007)

Other professional activities

- Program committee, North American Computational Linguistics Olympiad, (2008–2010, 2013–2018).
- Organizing committee, Twenty-Ninth and Thirtieth Annual Meetings of the Berkeley Linguistics Society (2002 and 2003).
- Special session organizer, Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society (2001). Co-organized with Patrick Chew.