

Xiaofei Lu

Department of Applied Linguistics
The Pennsylvania State University
305 Sparks Building
University Park, PA 16802, USA

814-865-4692 (phone)
814-865-7944 (fax)
xxl13 AT psu DOT edu
www.personal.psu.edu/xxl13

Research Interests

- *Corpus linguistics*: compilation, annotation, and analysis of native and learner corpora
- *Computational linguistics*: combining linguistic insight with machine learning methods for natural language processing
- *Intelligent computer-assisted language learning*: use of natural language processing technology for computer-assisted language learning

Education

- *Ph.D., Linguistics*, The Ohio State University, Columbus, Ohio, USA, 2006
Dissertation: *Hybrid Models for Chinese Unknown Word Resolution*
Committee: Detmar Meurers (advisor), Chris Brew and Marjorie Chan
- *M.A. (by research), English Language*, National University of Singapore, Singapore, 2002
Thesis: *The Taiwan Issue in the Media: A Corpus-Based Comparative Study of Chinese and American Media Discourse*
- *B.A., English*, Nankai University, Tianjin, China, 1997
Thesis: *A Thematic Study of Thomas Hardy's Four Major Novels*

Academic/Research Appointments

- *Assistant Professor of Applied Linguistics*, Department of Applied Linguistics, The Pennsylvania State University, 2006-present
- *Presidential Fellow*, The Ohio State University, 2005-2006
- *Graduate Research Intern*, Natural Language Theory and Technology Group, Palo Alto Research Center (PARC), Summer 2005
- *Graduate Teaching Associate*, Department of Linguistics, The Ohio State University, 2002-2005
- *University Fellow*, The Ohio State University, 2001-2002
- *Research Scholar*, National University of Singapore, 1999-2001

Publications

Edited volumes

Feldman, Anna and Lu, Xiaofei (eds.) (2007). *Proceedings of the NAACL HLT 2007 Workshop on Computational Approaches to Figurative Language*. Rochester, NY: The Association for Computational Linguistics.

Refereed journal articles

- Lu, Xiaofei and Thorne, Steven L. (in preparation). A historical analysis of linguistic changes in the American reading curriculum.
- Lu, Xiaofei (under review). The relationship of lexical richness to the quality of ESL learners' oral narratives. Revised and resubmitted to *The Modern Language Journal*.
- Lu, Xiaofei (under review). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language proficiency. Revised and resubmitted to *TESOL Quarterly*.
- Lu, Xiaofei (forthcoming). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*.
- Lu, Xiaofei (2009). Automatic measurement of syntactic complexity in child language acquisition. *International Journal of Corpus Linguistics*, 14(1):3-28.
- Lu, Xiaofei (2008). Improving part-of-speech guessing of Chinese unknown words using hybrid models. *International Journal of Corpus Linguistics*, 13(2):169-193.
- Lu, Xiaofei (2008). Hybrid models for sense guessing of Chinese unknown words. *International Journal of Corpus Linguistics*, 13(1):99-128.
- Lu, Xiaofei (2007). A hybrid model for Chinese word segmentation. *Journal for Computational Linguistics and Language Technology*, 22(1):71-88.

Book chapters

- Lu, Xiaofei (forthcoming). What can corpus software reveal about language development? In Anne O'Keeffe & Michael McCarthy (eds.), *Routledge Handbook of Corpus Linguistics*. Oxfordshire, UK: Routledge.
- Lu, Xiaofei (2002). Discourse and ideology: The Taiwan issue in the Chinese and American media. In Christopher N. Candlin (ed.), *Research and Practice in Professional Discourse*, pp. 589-608. Hong Kong: City University of Hong Kong Press.

Articles in conference proceedings

- Lu, Xiaofei (2008). Automatic measurement of syntactic complexity using the revised developmental level scale. In *Proceedings of the 21st International Florida Artificial Intelligence Research Society Conference (FLAIRS-08)*, pp. 153-158. Menlo Park, CA: AAAI Press.
- Lu, Xiaofei (2007). Combining machine learning with linguistic heuristics for Chinese word segmentation. In *Proceedings of the 20th International Florida Artificial Intelligence Research Society Conference (FLAIRS-07)*, pp. 241-246. Menlo Park, CA: AAAI Press.
- Lu, Xiaofei (2007). Hybrid models for semantic classification of Chinese unknown words. In *Proceedings of Human Language Technologies 2007: The Conference of the North American Chapter of the*

Association for Computational Linguistics (NAACL HLT 2007), pp. 188-195. Rochester, NY: The Association for Computational Linguistics. [Acceptance rate: 24%]

Lu Xiaofei (2005). Hybrid methods for POS guessing of Chinese unknown words. In *Proceedings of the Student Research Workshop at the 43rd Annual Meeting of the Association for Computational Linguistics* (ACL-05), pp. 1-6. Ann Arbor, MI: The Association for Computational Linguistics. [Acceptance rate: 11%]

Lu, Xiaofei (2005). Promoting learner awareness of language transfer errors in ICALL. In *Proceedings of the 2004 International Symposium on Applied Linguistics and Language Teaching* (ISALLT-04), pp. 521-530. Cedar Falls, IA: The University of Northern Iowa.

Lu, Xiaofei (2005). Towards a hybrid model for Chinese word segmentation. In *Proceedings of the 4th SIGHAN Workshop on Chinese Language Processing held at the 2nd International Joint Conference on Natural Language Processing* (IJCNLP-05), pp. 189-192. Jeju, Korea: Asian Federation of Natural Language Processing.

Lu, Xiaofei (2004). On the distribution of Chinese number expressions. In *Proceedings of the 20th Northwest Linguistics Conference* (NWLC-04), pp. 145-160. Seattle, WI: University of Washington.

Book reviews

Lu, Xiaofei (forthcoming). Review of Fletcher, Paul and Jon F. Miller (eds.) (2005). *Developmental Theory and Language Disorders* (Amsterdam: John Benjamins). *Word – Journal of the International Linguistic Association*.

Lu, Xiaofei (forthcoming). Review of Zubizarreta, Maria Luisa and Eunjeong Oh (2007). *On the Syntactic Composition of Manner and Motion* (Cambridge, MA: MIT Press). *Word – Journal of the International Linguistic Association*.

Lu, Xiaofei (forthcoming). Review of Xiao, Richard, Paul Rayson, and Tony McEnery (2009). *A Frequency Dictionary of Mandarin Chinese* (Oxfordshire, UK: Routledge). *Chinese Language and Discourse*, 1(1).

Lu, Xiaofei (2008). Review of Roark, Brian and Richard Sproat (2007). *Computational Approaches to Morphology and Syntax* (Oxford, UK: Oxford University Press). *Linguist List*, 19.438.

Lu, Xiaofei (2007). Review of de Klerk, Vivian (2006). *Corpus Linguistics and World Englishes: An Analysis of Xhosa English* (London, UK: Continuum). *Linguist List*, 18.1195.

Presentations

Lu, Xiaofei and Lantolf, James (2010). Tracking L2 development using a corpus-based approach. Paper to be presented at the 2010 Northeast Conference on the Teaching of Foreign Languages (NECTFL-10). New York, NY. March 25-27.

- Lu, Xiaofei and Thorne, Steven L. (2010). A historical analysis of linguistic changes in the American reading curriculum. Paper to be presented at the American Association for Applied Linguistics 2010 Conference (AAAL-10). Atlanta, GA. March 6-9.
- Lu, Xiaofei (2009). A corpus-based approach to tracking L2 development. Paper presented at the 2009 Annual Convention of the American Council on the Teaching of Foreign Languages (ACTFL-09). San Diego, CA. November 20-22.
- Lu, Xiaofei (2009). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language proficiency. Paper presented at the American Association for Applied Linguistics 2009 Conference (AAAL-09). Denver, CO. March 21-24.
- Heift, Trude, Schulze, Mathias, Meurers, Detmar, and Lu, Xiaofei (2009). Successes and challenges of ICALL for learning, teaching and research. Discussion panel presented at the 26th Annual Symposium of the Computer Assisted Language Instruction Consortium (CALICO-09). Tempe, AZ. March 10-14.
- Lu, Xiaofei (2009). Automating and evaluating measures of L2 lexical complexity. Paper presented at the 26th Annual Symposium of the Computer Assisted Language Instruction Consortium (CALICO-09). Tempe, AZ. March 10-14.
- Lu, Xiaofei (2009). Automating measures of L2 syntactic complexity. Paper presented at the Workshop on Automatic Analysis of Learner Language at the 26th Annual Symposium of the Computer Assisted Language Instruction Consortium (CALICO-09). Tempe, AZ. March 10-14.
- Lu, Xiaofei (2008). Corpus linguistics research at Penn State. Paper presented at the 2nd Meeting of the Pennsylvania Applied Linguistics Consortium (PAALC). Philadelphia, PA. September 27.
- Lu, Xiaofei (2008). Automatic measurement of syntactic complexity using the revised developmental level scale. Paper presented at the 21st International Florida Artificial Intelligence Research Society Conference (FLAIRS-08). Coconut Grove, FL. May 15-17.
- Lu, Xiaofei (2008). Automatic measurement of syntactic complexity in second language acquisition. Paper presented at the Workshop on Automatic Analysis of Learner Language at the 25th Annual Symposium of the Computer Assisted Language Instruction Consortium (CALICO-08). San Francisco, CA. March 18-22.
- Lu, Xiaofei (2007). Computational analysis of linguistic complexity. Paper presented at the 24th Annual Symposium of the Computer Assisted Language Instruction Consortium (CALICO-07). San Marcos, TX. May 22-26.
- Lu, Xiaofei (2007). Combining machine learning with linguistic heuristics for Chinese word segmentation. Paper presented at the 20th International Florida Artificial Intelligence Research Society Conference (FLAIRS-07). Key West, FL. May 7-9.
- Lu, Xiaofei (2007). Hybrid models for semantic classification of Chinese unknown words. Paper presented at Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics (NAACL HLT 2007). Rochester, NY. April 22-27.

- Lu, Xiaofei (2005). Knowledge representation matching. Poster presented at the Palo Alto Research Center (PARC) Summer Intern Poster Session. Palo Alto, CA. July 29.
- Lu, Xiaofei (2005). Hybrid methods for POS guessing of Chinese unknown words. Paper presented at the Student Research Workshop at the 43rd Annual Meeting of the Association for Computational Linguistics (ACL-05). Ann Arbor, MI. June 27.
- Lu, Xiaofei (2004). Are generics modal? Paper presented at the 34th Annual Meeting of the Michigan Linguistic Society (MLS-04). Flint, MI. October 16.
- Lu, Xiaofei (2004). Promoting learner awareness of language transfer errors in ICALL. Paper presented at the 6th International Symposium on Applied Linguistics and Language Teaching (ISALLT-04). Beijing-Shanghai, China. August 8-14.
- Lu, Xiaofei (2004). On the distribution of Chinese number expressions. Paper presented at the 20th Northwest Linguistics Conference (NWLC-04). Seattle, WA. May 1-2.
- Lu, Xiaofei (2000). Discourse and ideology: The Taiwan issue in the Chinese and American media. Paper presented at the International Conference on Research and Practice in Professional Discourse. Hong Kong, China. November 15-19.

Funded Research

External funding

- | | |
|-----------|--|
| 2008-2011 | Title: <i>Computerized Dynamic Assessment of Language Proficiency in French, Russian and Chinese</i>
Role: Consultant (Co-PIs: James P. Lantolf & Matthew E. Poehner)
Funded by: International Research and Studies Program, U.S. Department of Education
Award Amount: \$321,000 |
| 2008-2011 | Title: <i>Rereading Our Past: The Cognitive Demands of Reading and Reading Comprehension, 1890-2005</i>
Role: Investigator (PI: David A. Gamson)
Funded by: The Spencer Foundation
Award Amount: \$468,750 |
| 2006-2010 | Title: <i>Assessing Development of Advanced Proficiency through Learner Corpora</i> (one of 10 projects in CALPER)
Role: Project co-director (with Michael McCarthy)
Funded by: U.S. Department of Education (Title VI Award)
Award Amount: \$1,252,000 for all 10 projects in CALPER |

Student projects

- 2009 Title: *Using L2 Learner English to Improve English-Chinese Machine Translation*
 Role: Faculty advisor (Student advised: Chris Lengerich)
 Funded by: Undergraduate Summer Discovery Grant, PSU
 Award Amount: \$2,500
- 2007-2008 Title: *Analyzing the Real-World Strength between Geographically-Distinct Node Clusters on Facebook's Social Graph*
 Role: Faculty advisor (Student advised: Alan deLevie)
 Funded by: President's Fund for Undergraduate Research, PSU
 Award Amount: \$500

Teaching

- 2006-present Department of Applied Linguistics, The Pennsylvania State University
 (1) *APLNG 200: Introduction to Language, Culture and Social Interaction* (AU09)
 (2) *APLNG 597A: Seminar in Applied Corpus Linguistics* (AU09)
 (3) *APLNG 083S: Communication and the Internet* (AU08, AU07, AU06)
 (4) *APLNG 583: Language Testing* (SP08)
 (5) *APLNG 597C: Statistical Analysis of Qualitative and Corpus Data* (SP08)
 (6) *APLNG 588: Computer Assisted Language Learning* (AU07)
 (7) *APLNG 597D: Functional Linguistics* (SP07)
 (8) *APLNG 597E: Introduction to Corpus Linguistics* (SP07)
- Summer 2009 2009 Summer Institute in Applied Linguistics, University Park, PA
 (1) *APLNG 596D: Computational and Statistical Methods for Corpus Analysis*
- 2002-2005 Department of Linguistics, The Ohio State University
 (1) *LING 384: Language and Computers* (AU05, SP05, WI05, AU04)
 (2) *LING 201: Intro to Language in the Humanities* (SP04, WI04, AU03, SP03, AU02)
- 1997-1999 Department of Computer Science, Nankai University, Tianjin, China
 (1) *College English Intensive Reading* (SP99, AU98, SP98, AU97)
 (2) *English Grammar Refresher* (AU98)
- 1997-1999 Department of English, Xiehe University, Tianjin, China
 (1) *College English Composition* (SP99, AU98, SP98, AU97)
 (2) *College English Conversation* (SP99)

Honors**Fellowships and scholarships**

Presidential Fellowship, The Ohio State University, 2005-2006
Adjacent Technologies Fellowship, The Ohio State University, 2003
University Fellowship, The Ohio State University, 2001-2002

Research Scholarship, National University of Singapore, 1999-2001
First-Class Guanghua Scholarship, Guanghua Education Fund, China, 1998-1999
Baosteel National Education Scholarship, Baosteel Education Fund, China, 1996-1997
First-Class Nankai University Scholarship, 1994-1997

Awards

Undergraduate Teaching Award, Department of Linguistics, The Ohio State University, 2005
National Graduate School Entrance Examination Waiver, Nankai University, 1997
Outstanding Graduate of the Class of 1997 (top 1%), Nankai University, 1997
Outstanding Student of Nankai University, 1994-1997
Outstanding Student of Hunan Province, Hunan Provincial Education Commission, China, 1993

Travel grants

Student Travel Grant, The Association for Computational Linguistics, 2005
Four *Travel Grants*, Department of Linguistics, The Ohio State University, 2004-2005
Research Scholar Travel Grant, National University of Singapore, 2000

Memberships and Professional Service

Memberships

The Association for Computational Linguistics (ACL)
The Computer Assisted Language Instruction Consortium (CALICO)

Offices held in professional societies

Co-chair, Special Interest Group in Intelligent Computer-Assisted Language Learning of the Computer Assisted Language Instruction Consortium (ICALL SIG of CALICO), 2008-present

Conference organization

Co-chair, CALICO 2009 Preconference Workshop on Automatic Analysis of Learner Language, Tempe, AZ, 2009
Co-chair, NAACL HLT 2007 Workshop on Computational Approaches to Figurative Language, Rochester, NY, 2007
Co-chair, 2nd Midwest Computational Linguistics Colloquium, Columbus, OH, 2005

Reviewing

Editorial board member, *The Linguistics Journal*, 2007-present
Reviewer, *International Journal of Applied Linguistics*, 2009-present
Reviewer, *TESOL Quarterly*, 2006-present
Reviewer, *CALICO Journal*, Special Issue on Automatic Analysis of Learner Language, 2008
Reviewer, *Language Resources and Evaluation*, 2008
Reviewer, American Association for Applied Linguistics 2010 Conference, Atlanta, GA, 2010

Program Committee, NAACL HLT 2010 Workshop on Computational Approaches to Linguistic Creativity, Los Angeles, CA, 2010
Program Committee, Special Track on Applied Natural Language Processing, 23rd International Florida Artificial Intelligence Research Society Conference, Daytona Beach, FL, 2010
Program Committee, 13th Conference on Computational Natural Language Learning, Boulder, CO, 2009
Program Committee, NAACL HLT 2009 Workshop on Computational Approaches to Linguistic Creativity, Boulder, CO, 2009
Program Committee, Special Track on Applied Natural Language Processing, 22nd International Florida Artificial Intelligence Research Society Conference, Sanibel, FL, 2009
Program Committee, Special Track on Applied Natural Language Processing, 21st International Florida Artificial Intelligence Research Society Conference, Coconut Grove, FL, 2008
Program Committee, Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Vancouver, Canada, 2005
Reviewer, New Trends in Intelligent Information Processing and Web Mining, Gdansk, Poland, 2005
Reviewer, 20th International Conference on Computational Linguistics, Switzerland, Geneva, 2004
Reviewer, 11th Conference on Natural Language Processing, Fez, Morocco, 2004

Committee service

Senator, University Faculty Senate, PSU, 2008-present
Member, Senate Standing Committee on Faculty Benefits, PSU, 2008-present
Member, Liberal Arts Caucus of Senators, College of the Liberal Arts, PSU, 2008-present
Member, Curricular Affairs Committee, College of the Liberal Arts, PSU, 2007-present
Member, Graduate Admissions Committee, Department of Applied Linguistics, PSU, 2007-present
Member, ESL/ITA Director Search Committee, Department of Applied Linguistics, PSU, 2008-2009
Member, Awards Committee, Department of Applied Linguistics, PSU, 2006-2009
Judge, 2008 Undergraduate Exhibition, PSU, 2008
Judge, 22nd Annual Graduate Exhibition, PSU, 2007
Judge, 2nd Annual Chinese New Year Drama Night and Speech Contest, PSU, 2007
Presenter, Interdepartmental TA Preparation Workshop, Center for Language Acquisition, PSU, 2007
Member, The Ohio State University Working Papers in Linguistics Committee, 2003-2004
Treasurer, Student Linguistic Association, OSU, 2003-2004
Member, Travel Committee, Department of Linguistics, OSU, 2001-2003