

Na-Rae Han

MAILING ADDRESS:

Institute of Korean Culture
Seongbuk-Gu Anam-Dong 5-1
Seoul, 136-701
Republic of Korea

PHONE: +82-10-6262-******EMAIL:** naraehan@gmail.com**WEB:** <http://www.ling.upenn.edu/~nrh>

EDUCATION	Ph.D. in Linguistics University of Pennsylvania, Philadelphia, PA	May 2006
	<ul style="list-style-type: none">• Concentration: Computational Linguistics• Dissertation Title: <i>Korean Zero Pronouns: Analysis and Resolution</i>• Advisors: Ellen F. Prince, Martha S. Palmer	
	M.S.E. in Computer and Information Science University of Pennsylvania, Philadelphia, PA	May 2001
	M.A. in Linguistics Seoul National University, Seoul, Korea	July 1997
	B.A. in Anthropology Seoul National University, Seoul, Korea	February 1995

RESEARCH INTERESTS Natural language processing, corpus linguistics, NLP methods for educational assessment and instruction, reference resolution, computational morphology, pragmatics

RESEARCH EXPERIENCE	<i>Research Professor</i> Electronic Humanities Group Institute of Korean Culture Korea University, Seoul, Korea	November 2007–
	<ul style="list-style-type: none">• Conducting studies on large-scale Korean corpora, both raw and annotated, with the aim to discover distinct linguistic phenomena and investigate them in the larger context of present day Korean culture. The current goal is to identify and analyze linguistic features that characterize text genres, registers, gender, and authorship.	
	<i>Consultant</i> Automated Scoring and Natural Language Processing Group Educational Testing Service, Princeton, NJ	October 2006–December 2006
	<i>Associate Development Scientist</i> Automated Scoring and Natural Language Processing Group Educational Testing Service, Princeton, NJ	September 2005–September 2006
	<ul style="list-style-type: none">• Team leader of the <i>c-rater</i> project; <i>c-rater</i> provides content-based scoring capabilities for short constructed response tests by utilizing natural language processing techniques.• Developed applications that are aimed at aiding ESL (English as a Second Language) learners, such as linguistically engineered tools for correcting grammatical mistakes in essay writing.	
	<i>Research Consultant</i> Educational Testing Service/ETS Technologies, Inc., Princeton, NJ	Summer 2002, July–December 2003

- Worked in the Product Development and Research Application division, which focuses on developing computerized applications for modeling and scoring student test performances.

Research Assistant 2002-2003
Linguistic Data Consortium (LDC) at the University of Pennsylvania

- Developed transcription guidelines for the Korean Call-Friend corpus project, which produced a collection of transcribed Korean speech data; built an SQL database of Korean lexicon derived from the corpus
- Morphology Workbench project: built a corpus of morphologically annotated Korean texts.

Project Coordinator 2001-present
Project Member 1999-present
Penn Korean Treebank Project, University of Pennsylvania

- Developed annotation guidelines for part-of-speech tagging and syntactic bracketing for constructing annotated Korean texts; developed annotation tools; participated in the annotation process; evaluated the corpora for quality.

Project Member July-December 2001
Korean-English Machine Translation Project

- Joint project of University of Pennsylvania and Cogentex, Inc.
- Took part in developing a Korean-to-English machine translation system, which makes use of a set of linguistically motivated rules.

Visitor June–July 2000
Xerox Research Centre in Europe, Grenoble, France

- Developed Klex, a lexical dictionary lookup system for Korean, using Xerox’s Finite-State tools and development suites; final product was later released by LDC (2004).

SELECTED PUBLICATIONS

Martin Chodorow, Joel Tetreault and Na-Rae Han (2007) “Detection of Grammatical Errors Involving Prepositions” *Proceedings of the Fourth ACL-SIGSEM Workshop on Prepositions, Prague*.

Na-Rae Han, Martin Chodorow, and Claudia Leacock (2006) “Detecting Errors in English Article Usage by Non-Native Speakers,” *Natural Language Engineering: Special Issue on Educational Applications* 12(2), Cambridge University Press, UK.

Na-Rae Han, Shijong Ryu, Sook-Hee Chae, Seung-yun Yang, Seunghun Lee, and Martha Palmer (2006) “Korean Treebank Annotations Version 2.0” Linguistic Data Consortium (LDC) catalog number LDC2006T09 and ISBN 1-58563-381-X.

Na-Rae Han (2005) “Klex: A Finite-State Transducer Lexicon of Korean,” in *Proceedings of the 5th International Workshop on Finite-State Methods and Natural Language Processing*, Springer LNCS.

Na-Rae Han, Martin Chodorow, and Claudia Leacock (2004) “Detecting Errors in English Article Usage with a Maximum Entropy Classifier Trained on a Large, Diverse Corpus,” in *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon, Portugal.

Na-Rae Han (2004) “Korean Null Pronouns: Classification and Annotation,” in Bonnie Webber and Donna K. Byron, eds., *ACL 2004 Workshop on Discourse Annotation*, Barcelona, Spain, Association for Computational Linguistics, 33–40.

Na-Rae Han (2004) "Klex: Finite-State Lexical Transducer for Korean," Linguistic Data Consortium (LDC) catalog number LDC2004L01 and ISBN 1-58563-283-x.

Na-Rae Han (2004) "Morphologically Annotated Korean Text," Linguistic Data Consortium (LDC) catalog number LDC2004T03 and ISBN 1-58563-284-8.

Na-Rae Han, David Graff, and Myeonchul Kim (2003) "Korean Telephone Conversations Lexicon," Linguistic Data Consortium (LDC) catalog number LDC2003L02 and ISBN 1-58563-265-1.

Chung-hye Han, Na-Rae Han, Eon-Suk Ko and Martha Palmer (2002) "Development and evaluation of a Korean Treebank and its application to NLP," *Language and Information*, vol 6.1, pp 123-138.

Martha Palmer, Chung-Hye Han, Na-Rae Han, Eon-Suk Ko, Hee-Jong Yi, Alan Lee, Chris Walker, John Duda, and Nianwen Xue (2002) "Korean English Treebank Annotations," Linguistic Data Consortium (LDC) catalog number LDC2002T26 and ISBN 1-58563-236-8.

Chung-hye Han, Na-Rae Han, Eon-Suk Ko, and Martha Palmer (2002) "Development and Evaluation of a Korean Treebank and its Application to NLP," in *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*.

Chung-hye Han, and Na-Rae Han (2001) "Part of Speech Tagging Guidelines for Penn Korean Treebank," Technical report, IRCS, University of Pennsylvania.

Chung-hye Han, Na-Rae Han, and Eon-Suk Ko (2001) "Bracketing Guidelines for Penn Korean Treebank," Technical report, IRCS, University of Pennsylvania.

**TEACHING
EXPERIENCE**

Lecturer Spring 2008
Mathematical Background for Linguistics
Department of Linguistics, Seoul National University
A graduate-level course for Linguistics majors. This course covers mathematical foundations for linguistics, in particular semantics and formal language theories.

Lecturer Spring 2008
Topics in English Linguistics: Corpus Linguistics
Department of English Education, Hongik University
An elective course for English Education majors. This course is an introduction to the use of corpora in the study of language and language education.

Teaching Assistant Spring 2000
LING 105: Introduction to Cognitive Science, The University of Pennsylvania

Teaching Assistant Fall 1999
LING 106: Formal Methods in Linguistics, The University of Pennsylvania

**OTHER
EXPERIENCE**

TESOL Workshop Spring 2005
Participated in TESOL (Teacher of English to Speakers of Other Languages) Workshop, School of Education, University of Pennsylvania.

LANGUAGE SKILLS Korean: native
English: near-native
Japanese: proficient

COMPUTER SKILLS Languages: Proficient in perl, HTML, XML, Unix shell scripting; familiar with java
Operating Systems: Linux, Solaris, MS Windows, Cygwin.

HONORS AND AWARDS *IRCS Fellow* 2001-2002
awarded by Institute for Research in Cognitive Science, Univ. of Pennsylvania

Woosan Fellow 1995-1997
awarded by Woosan Foundation, Seoul Korea

SERVICE

- Committee: Workshop on Innovative Use of NLP for Building Educational Applications, Annual Meeting of ACL (2008)
- Reviewer: *Eoneohag* (“*Linguistics*”), Penn Linguistics Colloquium (2004-2007)
- Web Administrator: Department of Linguistics, the University of Pennsylvania (2001-2002)
- Coordinator: Speaker Series, Department of Linguistics, the University of Pennsylvania (2002)
- Editor: U. Penn Working Papers in Linguistics (1999)
- Committee: The 23rd Penn Linguistics Colloquium (1999)

REFERENCES Available Upon Request