

Corpus Linguistics for Korean Language Teachers
Graduate Program, MA in Korean
School of Korean, Middlebury College
Summer 2022

Instructor: Sun-Hee Lee (sunheel@middlebury.edu): she
Classes: TH 2:30-5:20 & F 2:30-4:20 5 hours per week
(Check the weekly schedule for activities)
Office Hours: Tuesday 4:30-5:30 & Friday 1:00-2:00

COURSE DESCRIPTION & GOAL:

This course presents a practical introduction to corpus linguistics, an extremely versatile methodology of language analysis for Korean language teachers. Advances in computer technology have revolutionized the ways linguists can approach their data. By accessing large digital bodies of text (corpora) and searching for phenomena of interest, we can uncover complexities in naturally occurring data and explore broader issues utilizing linguistic patterns and frequency information.

Some of the fundamental questions to explore are: what is a corpus, and what corpora exist? How are corpora constructed and linguistic annotation added? What tools are available for search, annotation, and analysis? In addition to acquiring theoretical knowledge and practical skills necessary to collect and analyze corpora, students will learn how corpora are used in diverse areas such as sociolinguistics, discourse analysis, and language acquisition as well as language learning and teaching and develop their own research ideas using corpus-based methodology and resources.

READING MATERIALS:

Reading assignments are specified on the syllabus, and these articles will be posted on the course site.

REQUIREMENTS

- Homework assignments: 20 %
- Midterm Project (Annotated Bibliography): 15 %
- Leading Discussion and Presentation: 15%
- Final Pilot Research Projects: 40 %
- Attendance and Class Participation: 10 %

1. Homework Assignments (20 %):

There will be in-class and homework assignments including comments and questions on the readings for every class. These assignments will give you an opportunity to practically explore the topics discussed in class. Assignments are due by the beginning of each class. You may post your comments and questions for readings in the CANVAS discussion. For individual assignments, you may submit it by email or by submitting them in class. You can work together on the homework assignments; however, you must write out your own answer.

2. Annotated Bibliography (15%): The midterm assignment will be an annotated bibliography that is connected to your pilot research topic. Each student will choose a topic of interest and collected the key readings. The deadline of submitting the annotated bibliography is before the class of Friday July 29. It is not allowed to work with other students for your midterm assignment.

3. Leading Discussion and Presentations (15%): Each student will be asked to lead one class discussion by choosing a given topic out of various case studies. You will also give a presentation after exploring some corpus resources and tools in class. Presentation files need to be submitted to the instructor after the presentation.

4. Final Research Project (Pilot Study) (40%): You will develop and conduct independent research on a topic related to corpora. For the final research project, you will need to conduct background research using class materials and, more importantly, other resources, in order to develop your own idea and to explore the topic. Each of you will have an individual meeting with the instructor and discuss research topics and plans. Progress reports on your project (a proposal in 500 words, a final presentation [20 minutes + 10minutes Q/A], and a paper in 2500 words excluding corpus data) must be submitted by the deadlines throughout the term. See the following schedule for details. The final paper must be submitted to the course assignment folder by 11:59 pm on August 19, 2022. Regulations pertaining to plagiarism and academic misconduct will be strictly observed. If you are unclear about them, please check with your instructor.

Notes:

- In thinking about your projects, note that you can start with general exploration of a corpus, but eventually you will have to pin down a research question that you are investigating, based on previous research (in linguistics, corpus linguistics, or a related discipline).

- Please use APA citation style and be sure that you are consistent in your use of APA in all written work you turn in.

5. Attendance & Class Participation (10%):

Each person's insights, questions, thoughts, etc., are extremely valuable. You are expected to contribute to the discussion on daily basis, and your level of participation will be reflected in your grade. Note that class attendance is not included in your grade per se, but if you are not in class, you cannot participate. Especially during the student presentation days, please make sure not to skip class since your feedback to the presenters will be highly valued and will also be counted for your class participation. Contributing extra comments in the course forum discussion will count for bonus credit.

- Your attendance and participation will be evaluated using the following rubric:

9–10 points	<ul style="list-style-type: none"> • Completes all tasks to best of ability by deadlines • Makes frequent and thoughtful contributions to class discussions • Consistently works to best of ability to demonstrate progress in learning the subject matter
7–8 points	<ul style="list-style-type: none"> • Completes most tasks at an acceptable level by deadlines • Makes thoughtful contributions to class, though could take advantage of more opportunities in discussion • Generally, works to best of ability to demonstrate progress in learning the subject matter
5–6 points	<ul style="list-style-type: none"> • Is frequently late in completion of tasks and fails to complete several tasks by deadlines • Contributes minimally to class discussions • Demonstrates a minimum effort towards progressing in the subject matter
4–1 points	<ul style="list-style-type: none"> • Does not complete most tasks • Does not contribute to class discussions • Does not demonstrate effort towards progressing in the subject matter

Other Notes on Grading or Late Submission: You will have two-time 48 hour “late passes” which you may turn in an assignment 48 hours past the set deadline without any point or grade deductions. If you want to use this 48-hour late pass, email me on the day in which the

assignment is due to alert me that you plan to turn in the assignment on the two days. The only exception to the 48-hour late pass rule is using it for the final project paper. If after you have used your two-time, 48-hour late passes exception and you still haven't turned in your work, 5% grade of the assigned work will be deducted per day. No credit will be given seven days from the due date. However, some extensions will be given to the COVID-19 related absences with a special notice from you and your program coordinator.

- *Assignments that are turned in late are not eligible for full credit –5% of the full credit will be subtracted per day and no credit will be given after seven days.*
- **Syllabus Modification**
Please be sure to regularly check the course site and your email for updates. I will also always post the most up-to-date copy of the syllabus on our course site.
- **Preferred Pronouns and Communication**
I will gladly honor your request to address you by your preferred name or pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. I also ask everyone to respect one another's pronoun use while remaining aware that patterns of speech and linguistic habits in a world beset by gender binaries can be difficult to change. My pronouns are typically she/her/hers.

Tentative Schedule

Week	Date	Topic	Readings & Homework
1	7/8	Basics & Types of Corpora	McEnery, T., Xiao, R., & Tono, Y. (2006). Unit A1 Corpus Linguistics: basics. McEnery, T. & Hardie, A (2012) ch.1
2	7/14	Understanding Corpus Design Corpus Annotation	Leech, G (1997) ch.1 Weisser, M. (2016) ch.3
	7/15	Corpora and Applied linguistics AntConc	McEnery, Ziao, & Tono (2006) A.10 Weisser, M. (2016) ch.9
3	7/21	AntConc & Collocations	Weisser (2016) ch.10 Gablasova, D., Brezina, V., & McEnery, T. (2017).
	7/22	Building a Corpus	Possible Research Topics (one paragraph) Due Alsop, S. and Nesi, H. (2009) Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017)
4	7/28	Learner Corpora & Annotation Leading Discussion 1	Research Proposal (500 words) Due Gilquin (2015) Gablasova, D., Brezina, V., McEnery, T. (2019)
	7/29	Converging Teaching and Corpora Leading Discussion 2	Gabrielatos (2005) Granger(2015)

5	8/4	Learner Corpus Annotation Topics: Lexical Bundle Formulaic Expressions Leading Discussion 3	Midterm Annotated Bibliography Due van Rooy B. (2015) Byber, D. (2009) Byber D., Conrad, S. & Cortes, V (2004)
	8/5	Topics: Stance Expressions Leading Discussion 4	Gablasova, D., Brezina, V., McEnery, T., & Boyd, E. (2017)
6	8/11	Resources for Learner Corpus Research Resource Presentations	TBA
	8/12	Resources for Learner Corpus Research (Korean learner corpora)	Lee, S., Dickinson, M., & Israel, R. (2016) & Lee (2022)
7	8/18	Future of Learner Corpus Research	McEnery, T., Brezina, V., Gablasova, D. & Banerjee, J.(2019)
	8/19	Final Student Presentations	

*****Disclaimer: This syllabus is subject to change. The most up-to-date version of this syllabus can be found on the course site. Please check the SYLLABUS folder in the course site regularly for assignment and readings.

READINGS

Week	Date	Readings
1	7/8 Friday	McEnery, T., Xiao, R., & Tono, Y. (2006). <i>Corpus-Based Language Studies: an advanced resource book</i> . Unit A1 Corpus Linguistics: basics. Available at https://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xCBLS/CBLS.htm McEnery, T. & Hardie, A (2012). <i>Corpus Linguistics: Method, Theory and Practice</i> , Chapter1 . pp.1-24. Cambridge University Press, Cambridge.
2	7/14	Leech, G. (1997) <i>Introducing corpus annotation</i> . (Eds. by) Garside, R., Leech, G. and McEnery, T. <i>Corpus Annotation</i> . 1-18. Longman. Weisser, M. (2016). Chapter.3 <i>Practical Corpus Linguistics</i> . 29-42. WILEY Blackwell.
	7/15	McEnery, Ziao, & Tono (2006) Unit A10. Corpora and Applied Linguistics. <i>Corpus-Based Language Studies: an Advanced Resource Book</i> . 80-122. Routledge. Available at https://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xCBLS/CBLS.ht Weisser, M. (2016). Chapter 9. Basic frequency analysis, <i>Practical Corpus Linguistics</i> . 146-179. WILEY Blackwell.
3	7/21	Weisser, M. (2016). <i>Practical Corpus Linguistics</i> . Chapter 10. Exploring words in context. 193-210. WILEY Blackwell. Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. <i>Language Learning</i> , 67 (S1), 155–179.
	7/22	Alsop, S. & Nesi, H. (2009). Issues in the development of the British academic written English (BAWE) corpus, <i>Corpora</i> , Volume 4, Issue 1, 71-83. Davies, M. (2012). Expanding horizons in historical linguistics with the 400-Million word corpus of historical American English, <i>Corpora</i> , Volume 7, Issue 2, 121-157. Optional: Love, R., Dembry, C., Hardie, A., Brezina, V.,& McEnery, T. (2017). The spoken BNC2014: designing and building a spoken corpus of everyday conversations. <i>International Journal of Corpus Linguistics</i> , 22:3.

4	7/28	<p>Gilquin, G. (2015) From design to collection of learner corpora. Granger, S., Gilquin, G., & Meunier, F. (Eds.) <i>The Cambridge Handbook of Learner Corpus Research</i>. 9-34. Cambridge University Press.</p> <p>Gablasova, D., Brezina, V., & McEnery, T. (2019) The Trinity Lancaster corpus: development, description and application. <i>International Journal of Learner Corpus Research</i>. 5. 126-158</p>
	7/29	<p>Gabrielatos, (2005) Corpora and language teaching: Just a fling or wedding bells? <i>The Electronic Journal for English as a Second Language (TEJE-EJ)</i> Vol. 8.4.</p> <p>Granger, S. (2015) The contribution of learner corpora to reference and instructional materials design. Granger, S., Gilquin, G., & Meunier, F. (Eds.) <i>The Cambridge Handbook of Learner Corpus Research</i>. Pp. 485-510. Cambridge University Press.</p>
5	8/4	<p>van Rooy B. (2015) Annotating learner corpora. Granger, S., Gilquin, G., & Meunier, F. (Eds.) <i>The Cambridge Handbook of Learner Corpus Research</i>. Pp. 79--105. Cambridge University Press.</p> <p>Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical Bundles in University Teaching and Textbooks. <i>Applied Linguistics</i>, 25, 371-405.</p> <p>Optional: Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. <i>International Journal of Corpus Linguistics</i>, 14(3), 275-311.</p>
	8/5	<p>Gablasova, D., Brezina, V., McEnery, T., & Boyd, E. (2017). Epistemic stance in spoken L2 English: the effect of task and speaker style. <i>Applied Linguistics</i>, 38(5), 613–637.</p>
6	8/11	TBA
	8/12	<p>Lee, S., Dickinson, M., & Israel R. Challenges of learner corpus annotation: focusing on Korean Learner Language Analysis (KoLLA) system. <i>Language Facts and Perspectives</i>, 38, 221-251.</p> <p>Lee, S. (2022). Corpus-based research and KSL, Byon, A.S., & Pyun, D.O. (Eds.). (2022). <i>The Routledge Handbook of Korean as a Second Language</i> (1st ed.). Routledge. https://doi.org/10.4324/9781003034704.</p>
7	8/18	<p>McEnery, T., Brezina, V., Gablasova, D., & Banerjee, J. (2019). Corpus linguistics, learner corpora, and SLA: Employing technology to analyze language use. <i>Annual Review of Applied Linguistics</i>, 39, 74-92.</p>
	8/19	Student Presentations