

LING 476/576: CORPUS LINGUISTICS

T/Th 12-1:50 – Tues. in Parkmill 52 classroom and Thur. in FMH B157 computer lab

Professor: Susan Conrad

E-mail: conrads@pdx.edu

Office: 335E UCB

Phone: 725-8727

Drop-in Office Hours: Tues. 2-2:45 & Wed. 11:30-12:30. You can come during those times without an appointment. E-mail for an appointment at another time.

There are no required prerequisites for this course. However, familiarity with grammar and morphology, such as covered in the Structure of English (Ling 392) or Applied English Grammar (Ling 521), is useful. The course meets the following requirements in programs in the Department of Applied Linguistics:

- MA TESOL program requirements starting Fall 2019: Language Use course
- MA TESOL program requirements before Fall 2019: Language Education or Language and Society credit. Talk with me about your choice.
- Certificate in TESL before Fall 2019: Language Education elective or unspecified elective
- B.A. in Applied Linguistics: Elective
- Minor in Applied Linguistics: Elective

COURSE DESCRIPTION AND OBJECTIVES

Corpus linguistics is an approach to analyzing language that uses large electronic collections of texts and computer-assisted analysis techniques. In this course, you will learn how to use corpus linguistics to investigate words and grammar, and how to apply the findings to language teaching or other language-related issues in the world.

One of the major advantages of corpus linguistics is that it gives us information about how language is actually used, rather than just focusing on what is grammatical/ungrammatical or how we *think* language is used. Many investigations of the interaction of language and society benefit from corpus techniques – e.g., critical discourse analysis, world Englishes, register analysis, dialect studies, English as a Lingua Franca, historical linguistics, forensic linguistics, and many more. Principles and techniques from corpus linguistics are also useful within the field of computational linguistics. Second language learners and teachers like corpus linguistics because it makes it possible for them to investigate language for themselves instead of relying solely on published materials or a native speaker's intuition. Corpus linguistics also provides some teaching options that are especially useful for developing learner autonomy.

The major focus of the course will be on *doing* corpus linguistics, not just reading about it. Specifically, the course will help you learn to do the following:

1. Design appropriate investigations for corpus-based analyses, including feasible questions to ask, appropriate corpora to use, and effective techniques to employ.
2. Conduct small corpus analysis projects, focusing on grammar, words, or lexico-grammar. (Some of these features will have discourse-level functions, but doing a corpus-based analysis of a discourse-level feature requires a more advanced course in corpus linguistics.)
3. Use concordancing software and online corpus sites effectively and efficiently.

4. Develop and discuss arguments for the appropriate use of corpus linguistics, and explain those situations or questions which are not suitable for corpus-based analyses.
5. Develop and conduct a substantial corpus-based research project or design and propose a principled, feasible new corpus. (If you are in the “old” curriculum of the TESL Certificate or MA TESOL program and are taking this course as a Language Education elective, you also have the option of a materials development project.)

See additional information about program learning objectives at the end of this document.

MATERIALS

- Course readings available to download from the library, an internet site, or our course D2L site.
- On loan to you during the course: a USB flash drive with a term-long version of the software MonoConc Pro 2.2 and the free Antconc, plus a copy of several corpora. The files must not be copied and the flash drive must be returned at the end of the quarter to receive a grade for the course.
- Required additional materials: a system of some sort for storing instructions, handouts, notes, etc., in an organized way. It’s especially important that you have easy access to earlier handouts and notes as you work on analyses. You can use electronic or hard copies, but your system has to be accessible, efficient and effective for you.

REQUIREMENTS, RESPONSIBILITIES AND GRADING

	points	
1. Assignments and Grades	<u>476</u>	<u>576</u>
Class preparation & participation (includes attendance & in-class work)	5	5
Worksheets (computer lab activities)	30	25
Analysis Paper	20	15
Project (larger for graduate students)	40	40
Feedback to other students about projects	5	5
Summary and Reaction Paper (for two published articles)	<u>n/a</u>	<u>10</u>
Total	100	100

Detailed information about the requirements for assignments will be posted on D2L. Read the information carefully and ask any questions you have about the requirements. Some assignments have multiple options. Choose the option that best fits your interests, or talk with me if you have trouble choosing.

And, remember, a B or better is passing for the M.A. TESOL program and a C or better for the TESL Certificate and B.A. in Applied Linguistics. A grade of B starts at 83/100 and a C starts at 73/100.

2. Late papers: Points will be deducted for late assignments unless you arrange for an extension at least **two weekdays before** the assignment is due. If you know you will have a problem finishing some work on time, talk with me about it. I do not give extensions automatically, but I do realize that sometimes parts of your life besides corpus linguistics need attention.

3. Readings: I have deliberately kept the readings for this course limited, mostly because I want more of your time to go into doing analyses rather than reading about them. However, all the readings are journal articles or book chapters, rather than a textbook, and some might be challenging. Use the study guides that are on D2L; they will tell you where to focus your attention. There is a list of additional readings on D2L if you want to explore other topics or go into more depth, or feel free to ask for suggestions.

4. Checking D2L: Check the announcements on our D2L site at least twice each week – on Wednesdays and sometime between Fri. and Monday – for any reminders for the next class. Checking more often is better.

5. General responsibilities as a member of this class: This is a 400/500-level course, and I have certain expectations of you as upper-level and graduate students: take responsibility and initiative in completing all work on time; be prepared for each class, including doing the readings or assigned analyses before class; participate in all class activities; work cooperatively and actively with other students; seek help when you have questions or problems; don't miss class – or if you absolutely must, ask if you can make up work you will miss (some can be made up; some can't) and afterwards get notes from other students. E-mail or talk with me in advance if you know you have to miss a class for an unavoidable conflict on a certain day. If you miss a class for an emergency, e-mail me when you can to let me know why you were absent and when you will be back.

Consistent with PSU's code of conduct requirements, I also expect behavior that is consistent with making PSU a safe and respectful environment for learning. This principle applies to class time and group work outside of class. This principle covers general behavior, such as treating all individuals respectfully regardless of race, religion, gender identity, or other characteristics, and also classroom behavior, such as helping to create a positive learning environment by staying focused on your learning and not on text messages, social media or other distractions.

If you ever perceive our class as not conducive to your learning, please come and talk with me about your concerns.

6. Academic conduct for individual work and group work: In class, some work will be done in pairs or groups. You can also work on assignments together outside of class unless I specifically ask you not to. In fact, I recommend working together because it is often useful both for solving technology problems and for interpreting language use. However, anything you hand in should be a result of your active participation, be written in your own words, and reflect your understanding. If you use others' words or ideas, be sure to quote and provide the reference. Consistent with PSU's Student Code of Conduct, I will give you an F or O for any assignment in which academic misconduct occurred. I will also report your academic misconduct to the Dean of Students office, which keeps track of individuals' academic misconduct throughout the university. Here is PSU's definition of academic misconduct (<https://www.pdx.edu/dos/psu-student-codeconduct#GeneralPolicy>):

Academic Misconduct is defined as, actual or attempted, fraud, deceit, or unauthorized use of materials prohibited or inappropriate in the context of the academic assignment. Unless otherwise specified by the faculty member, all submissions, whether in draft or final form, must either be the Student's own work, or must clearly acknowledge the

source(s). Academic Misconduct includes, but is not limited to: (a) cheating, (b) fraud, (c) plagiarism, such as word for word copying, using borrowed words or phrases from original text into new patterns without attribution, or paraphrasing another writer's ideas; (d) the buying or selling of all or any portion of course assignments and research papers; (e) performing academic assignments (including tests and examinations) in another person's stead; (f) unauthorized disclosure or receipt of academic information; (g) falsification of research data (h) unauthorized collaboration; (i) using the same paper or data for several assignments or courses without proper documentation; (j) unauthorized alteration of student records; and (k) academic sabotage, including destroying or obstructing another student's work.

If you feel unsure about what constitutes plagiarism or other academic misconduct, please come and talk with me about it.

7. Diversity, Inclusion, and Academic Accommodations: If you are registered with the Disability Resource Center (DRC) and need academic accommodations, please talk or e-mail with me about the message I received from the DRC about your accommodations.

Please be aware that any accessible tables or chairs in the room should remain available for students who find that standard classroom seating is not useable.

PSU values diversity and inclusion; we are committed to fostering mutual respect and full participation for all students. My goal is to create a learning environment that is equitable, inclusive, and as effective as possible for all students. If you experience any aspects of our course as barriers to your inclusion or learning, please notify me in person or via email. The Disability Resource Center (DRC) provides reasonable accommodations for students who encounter barriers in the learning environment.

8. Reporting of Sexual Discrimination, Harassment, and Violence: Portland State is committed to providing an environment free of all forms of prohibited discrimination and sexual harassment (sexual assault, domestic and dating violence, and gender or sex-based harassment and stalking). If you have experienced any form of gender or sex-based discrimination or sexual harassment, know that help and support are available. PSU has staff members trained to support survivors in navigating campus life, accessing health and counseling services, providing academic and on-housing accommodations, helping with legal protective orders, and more. Information about PSU's support services on campus, including confidential services and reporting options, can be found on PSU's Sexual Misconduct Prevention and Response website at: <http://www.pdx.edu/sexual-assault/get-help> or you may call a confidential IPV Advocate at 503-725-5672. You may report any incident of discrimination or discriminatory harassment, including sexual harassment, to either the Office of Equity and Compliance or the Office of the Dean of Student Life.

Please be aware that all PSU faculty members and instructors are required to report information of an incident that may constitute prohibited discrimination, including sexual harassment and sexual violence. This means that if you tell me about a situation of sexual harassment or sexual violence that may have violated university policy or student code of conduct, I have to share the information with my supervisor, the University's Title IX Coordinator or the Office of the Dean of Student Life. For more information about Title IX please complete the required student module Creating a Safe Campus in your D2L.

9. PSU Emergency Preparedness: For information about PSU emergency preparedness, please go to the Fire and Life Safety webpage (<https://www.pdx.edu/environmental-health-safety/fire-and-life-safety>).

TENTATIVE SCHEDULE

I may adjust the schedule based on student interests and needs. You are responsible for knowing any schedule changes that are announced in class even if you are late or absent. Changes will also be posted on D2L.

In addition to the readings, there will be homework doing corpus analyses. Plan to spend time preparing for every class.

Wk	Date	Class Topic or New Skills in Lab	Readings or Other Work Due (do the reading before the class)
1	Tues. Jan. 7	Introduction to the course and overview of corpus linguistics	Read syllabus thoroughly after class
	Thur. Jan. 9	Lab: Introduction to concordancing with MonoConc – KWIC files, word lists, collocates	Bennett 2010 (parts 1 and 2) Conrad 2011
2	Tues. Jan. 14	Lexical studies – words, multi-word sequences (bundles and clusters)	1. All read: Stubbs 1996 (chap 4) 2. Choose one of these: <ul style="list-style-type: none"> • Khani & Tazek, 2013 - if you are more interested in teaching/EAP or • Forchini 2013 – if you are more interested in translation or linguistic study generally (not teaching)
	Thur. Jan. 16	Lab: More with lexical searches, especially multiword sequences (lexical bundles, N-grams, clusters)	Week 1 lab worksheet due by today If you do not have a login for COCA, watch Kylie Moore's video 1 (especially the 2 nd ½) and create an account. Make sure you know your COCA password to use during lab. If you are not familiar with KWIC and Chart searches with COCA, watch Kylie Moore's video 2.
3	Tues. Jan. 21	Keyword analysis (not the same as Keyword in Context)	Culpepper 2009
	Thur. Jan. 23	Lab: Keyword analysis	Week 2 lab worksheet due by today

4	Tues. Jan. 28	Grammar and lexico-grammar	Conrad 2010 Biber, Gray & Poonpon 2011
	Thur. Jan. 30	Lab: Investigating grammar through text searches and regular expression searches	Week 3 lab worksheet due by today Analysis Paper due
5	Tues. Feb. 4	Corpus linguistics and critical discourse analysis	All read: Baker and Levon 2015 Required for grad, optional for undergrad: Ray 2016
	Thur. Feb. 6	Lab: Using tagged texts	Week 4 lab worksheet due by today
6	Tues. Feb. 11	Corpus linguistics and language instruction: Data-driven learning, analysis of textbooks During class: Look at Conrad & Biber 2009 unit 2, CEWP units 1 and 5.	Huang 2014 Vincent & Nesi 2018
	Thur. Feb. 13	Lab: More with checking results and making interpretations	Week 5 lab worksheet due by today Summary & reaction paper (grad students only) – due by Fri. Feb. 14 by 5 pm
7	Tues. Feb. 18	Corpus Design and Compilation	Reppen 2010 Clancy 2010
	Thur. Feb. 20	Lab: Compile and search your own mini-corpus	Week 6 lab worksheet due by today Bring texts to make a mini-corpus in lab Statement of project topic and focus due by today
8	Tues. Feb. 25	World Englishes and English as a Lingua Franca investigations	Leimgruber 2016 Chen 2017
	Thur. Feb. 27	Lab: GloWbE searches. Look at ACE, VOICE, ELFA sites.	Week 7 lab worksheet due by today
9	Tues. Mar. 3	Topic: To Be Determined (Learner corpora? Multi-dimensional analysis?)	To Be Determined
	Thur. Mar. 5	Lab: To Be Determined + 1 hour project work time	Week 8 lab worksheet due by today Bring draft of project in progress to work on during lab time
10	Tues. Mar. 10	Review and looking ahead in corpus linguistics	Tues.: Project due at the start of class. Projects exchanged with colleagues during class. Between Tues. and Thur.: Spend at least 2-3 hours reviewing a colleague's project

Thur. Mar. 12	Lab: Exchange project feedback. Write final assessment of your project.	Feedback to colleague due at the start of class.
Final Exam Time Thur. March 19 10:15-12:05	Save this time for a class meeting. There is no exam, but we will have a final project-sharing and wrap-up meeting.	

Course Readings (in order of use)

Instructions

- Unless otherwise noted, the items are available through an electronic resource link via the PSU library. Start at the PSU library website. Search for the **journal title**, follow links for the **online resource**, and when you make it to the journal website, look for the correct volume and pages in the archives or past issues list.
- If the entry says "See D2L," check our D2L site to get the reading.

Bennett, G. (2010). Introduction and chapter 1: Principles of corpus linguistics. *Using corpora in the language learning classroom*. Ann Arbor, MI: University of Michigan Press. [See D2L]

Conrad, S. (2011). Variation in corpora and its pedagogical implications. In V. Viana, S. Zyngier & G. Barnbrook (Eds.), *Perspectives on corpus linguistics* (pp. 47-62). Amsterdam: John Benjamins. [See D2L]

Stubbs, M. (1996). Chapter 4: Baden-Powell: A comparison of two short texts. *Text and corpus analysis*. Oxford, UK: Blackwell. [See D2L]

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC Journal*, 44(2), 209-232.

Forchini, P. (2013). A diachronic study of familiarizers ('man', 'guys', 'buddy', 'dude') in movie language. *Perspectives: Studies in Translatology*, 21(4), 504-525. [See D2L.]

Moore, K. (2017). COCA video tutorials for ESL teachers. (Master's project). Portland State University, Portland, Oregon. [Link in Additional Resources folder on D2L]

Culpeper, J. (2009). Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's Romeo and Juliet. *International Journal of Corpus Linguistics*, 14 (1), 29-59.

Conrad, S. (2010). What can a corpus tell us about grammar? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 227-240). Abingdon, UK: Routledge. [See D2L]

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35.

Baker, P., & Levon, E. (2015). Picking the right cherries? A comparison of a corpus-based and qualitative analysis of news articles about masculinity. *Discourse & Communication*, 9(2), 221-236.

- Ray, B. (2016). Stylizing genderlect online for social action: A corpus analysis of 'BIC Cristal for Her' reviews. *Written Communication*, 33(1), 42-67.
- Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, 26(2), 163-183.
- Vincent, B., & Nesi, H. (2018). The BAWE quicklinks project: A new DDL resource for university students. *Lidil* [Online], 58. <http://journals.openedition.org/lidil/5306>
- Conrad, S., & Biber, D. (2009). *Real grammar: A corpus-based approach to English*. White Plains, NY: Pearson. [See D2L. We will look at Unit 2 – He's looking at me... Progressive verbs vs. simple verbs.]
- CEWP – The Civil Engineering Writing Project: www.cewriting.org. Click on "Materials for Courses and Self-Study." We will look at parts of Language Unit 1 and Language Unit 5.
- Leimgruber, J. (2016). *Bah* in Singapore English. *World Englishes*, 35(1), 79-97.
- Chen, W. (2017). Functions of *so* in English as a lingua franca and native speech: A corpus-based study. *Chinese Journal of Applied Linguistics*, 40(1), 93-111. [Download for free directly from the journal website: <https://www.degruyter.com/view/j/cjal>]
- Reppen, R. (2010). Building a corpus. What are the key considerations? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 31-37). Abingdon, UK: Routledge. [See D2L]
- Clancy, B. (2010). Building a corpus to represent a variety of language. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 80-92). Abingdon, UK: Routledge. [See D2L]
-

Additional Information about Learning Outcomes

MA TESOL Program Learning Outcomes in the Corpus Linguistics Course

Students will be able to...

Language Use

1. explain how context, genre, register, and other factors influence variation in language.
2. describe the potential consequences of changing or not changing one's use of language based on the factors above.
3. describe the relationship between language structure and language use for a variety of target texts and genres.
4. identify specific linguistic structures and functions in samples of language in use.
5. articulate how language is used to create meaning in context, for example to convey information about social roles and relationships or speaker intents and beliefs.
6. describe some of the multiple facets of language competence, especially having to do with pattern vs. creativity.

Professional Skills and Dispositions

7. read and summarize primary literature in language teaching and applied linguistics.
8. produce an academic paper in which they support analytic claims with their data analysis using the writing conventions of the field.

9. make effective and appropriate use of technology for research, teaching, or other applied linguistics pursuits.
10. communicate and collaborate effectively and appropriately with colleagues and supervisors.
11. evaluate feedback and incorporate insights from it into their professional practice.
12. recognize English language users in all countries and at all levels as legitimate language users and question native speaker bias.
13. view language users as having agency and being able to make choices about their language use and learning.

Additional Optional Learning Outcomes – these depend on how you choose to focus your class assignments

1. analyze language teaching textbooks, materials, and technology for appropriateness, pedagogical approach, and fit with learner needs.
2. articulate the connection between theories and research in language teaching, and the connection between research and teaching practice.
3. determine the relevance of research findings for specific teaching and learning contexts.
4. evaluate the language of learners for their appropriateness and effectiveness.
5. analyze learner language to determine developmental trajectories and patterns of variation.

BA in Applied Linguistics Program Goals

The Corpus Linguistics course meets primarily the following goals:

- 1. Analysis Skills:** Students will exit the program with skills in analyzing language data and building arguments with language data, including basic skills for analyzing language data, using techniques from functional traditions, and developing, supporting and presenting written and spoken arguments involving language data.
- 2. Understanding of Theories:** Students will exit the program with a solid understanding of the role and nature of theories in the field of Applied Linguistics, including an awareness of functional approaches to language.
- 3. Application of Knowledge to the World:** Students will exit the program with the ability to apply their linguistic knowledge and analytical skills to a variety of language-related problems in the world, including the ability to speak and write about language data or language issues for a variety of audiences (e.g., explaining to a non-linguist how language choices affect the message that is conveyed).
- 4. Understanding of the Affective Nature of Language:** Students will exit the program with a conscious awareness of their own and other people's emotional reactions to language, language issues, and linguistic diversity, including the ability to describe the role of language in creating individual and group identities.
- 5. Academic and Professional Skills:** Students will exit the program with a skill set that enhances their ability to gain employment or continue on to graduate school, including effective use of technology and the ability to follow the style conventions for technical writing in applied linguistics.