1. General Information

1a. Submitted by the College of: ARTS &SCIENCES

Date Submitted: 12/8/2015

1b. Department/Division: Linguistics

1c. Contact Person

Name: Rusty Barrett

Email: erbarr2@uky.edu

Phone: 257-3114

Responsible Faculty ID (if different from Contact)

Name: Mark Richard Lauersdorf

Email: lauersdorf@uky.edu

Phone: 257-7101

1d. Requested Effective Date: Semester following approval

1e. Should this course be a UK Core Course? No

2. Designation and Description of Proposed Course

2a. Will this course also be offered through Distance Learning?: No

2b. Prefix and Number: LIN 540

2c. Full Title: Laboratory in Linguistics (subtitle required)

2d. Transcript Title: Lab in Linguistics

2e. Cross-listing:

2f. Meeting Patterns

LABORATORY: 2

2g. Grading System: Letter (A, B, C, etc.)

2h. Number of credit hours: 1

2i. Is this course repeatable for additional credit? Yes

If Yes: Maximum number of credit hours: 10

If Yes: Will this course allow multiple registrations during the same semester? Yes

2j. Course Description for Bulletin: A laboratory course offering students the opportunity for hands-on application of the general theories and methods of linguistics at the level of advanced undergraduate/beginning graduate training. The lab environment will generally involve both individual and small group work, developing both independent research skills and an ability to engage in collaborative linguistic investigation. May be repeated for credit under different topics.

2k. Prerequisites, if any: LIN 221 or consent of instructor; may require concurrent enrollment in an accompanying LIN lecture course.

2l. Supplementary Teaching Component:

3. Will this course taught off campus? No

If YES, enter the off campus address:

4. Frequency of Course Offering: Spring,

Will the course be offered every year?: Yes

If No, explain:

5. Are facilities and personnel necessary for the proposed new course available?: Yes

If No, explain:

6. What enrollment (per section per semester) may reasonably be expected?: 25

7. Anticipated Student Demand

Will this course serve students primarily within the degree program?: Yes

Will it be of interest to a significant number of students outside the degree pgm?: No

If Yes, explain:

8. Check the category most applicable to this course: Not Yet Found in Many (or Any) Other Universities ,

If No, explain:

9. Course Relationship to Program(s).

a. Is this course part of a proposed new program?: Yes

If YES, name the proposed new program: PhD in Linguistics

b. Will this course be a new requirement for ANY program?: No

If YES, list affected programs:

10. Information to be Placed on Syllabus.

a. Is the course 400G or 500?: Yes

b. The syllabus, including course description, student learning outcomes, and grading policies (and 400G-/500-level grading differentiation if applicable, from **10.a** above) are attached: Yes

Distance Learning Form

Instructor Name:

Instructor Email:

Internet/Web-based: No

Interactive Video: No

Hybrid: No

1.How does this course provide for timely and appropriate interaction between students and faculty and among students? Does the course syllabus conform to University Senate Syllabus Guidelines, specifically the Distance Learning Considerations?

2.How do you ensure that the experience for a DL student is comparable to that of a classroom-based student's experience? Aspects to explore: textbooks, course goals, assessment of student learning outcomes, etc.

3.How is the integrity of student work ensured? Please speak to aspects such as password-protected course portals, proctors for exams at interactive video sites; academic offense policy; etc.

4.Will offering this course via DL result in at least 25% or at least 50% (based on total credit hours required for completion) of a degree program being offered via any form of DL, as defined above?

If yes, which percentage, and which program(s)?

5.How are students taking the course via DL assured of equivalent access to student services, similar to that of a student taking the class in a traditional classroom setting?

6.How do course requirements ensure that students make appropriate use of learning resources?

7.Please explain specifically how access is provided to laboratories, facilities, and equipment appropriate to the course or program.

8.How are students informed of procedures for resolving technical complaints? Does the syllabus list the entities available to offer technical help with the delivery and/or receipt of the course, such as the Information Technology Customer Service Center (http://www.uky.edu/UKIT/)?

9.Will the course be delivered via services available through the Distance Learning Program (DLP) and the Academic Technology Group (ATL)? NO

If no, explain how student enrolled in DL courses are able to use the technology employed, as well as how students will be provided with assistance in using said technology.

10.Does the syllabus contain all the required components? NO

11.I, the instructor of record, have read and understood all of the university-level statements regarding DL.

Instructor Name:

LIN 540
Laboratory in Linguistics
Required subtitle: Corpus Linguistics

Instructor: *Mark Richard Lauersdorf* **Email:** *lauersdorf@uky.edu*
Meeting time: *TBD* **Classroom:** TBD
Office phone: *859---257---7101* **Office address:** *1471 POT*
Preferred method of contact: *email* **Office Hours:** *MWF 4:00–5:00 pm*

Prerequisite: LIN 221 or consent of instructor; may require concurrent enrollment in an accompanying LIN lecture course.

Course description: A laboratory course offering students the opportunity for hands---on application of the general theories and methods of linguistics at the level of advanced undergraduate/beginning graduate training. The lab environment will generally involve both individual and small group work, developing both independent research skills and an ability to engage in collaborative linguistic investigation. May be repeated to a maximum of ten credits.

Course objectives: This version of LIN 540 accompanies the lecture course LIN 510 Corpus Linguistics. We will focus on digital and computational tools and methods for: the construction of linguistic corpora, the use of corpora for linguistic investigations, and the analysis of corpus data resulting from those investigations.

Student learning outcomes: Upon completion of this lab students will be able to:

- use a variety of digital and computational tools for gathering and processing language data to build a linguistic corpus;
- implement various encoding schemas for annotation of linguistic corpus data and evaluate their appropriateness for different corpus designs;
- deploy different automated annotation tools and assess their accuracy on the basis of both precision and recall testing;
- determine the most appropriate data---mining tools and methods for different types of corpus---based linguistic investigations and analyze their effectiveness;
- carry out corpus---based linguistic investigations using both un---annotated and annotated corpora;
- perform linguistic analyses of corpus data including using statistical and visual methods to carry out data---driven analysis.

Required materials: In this lab we will be working with the following pieces of software, most of which are Open Source or available through campus licensing or as a web service (other software may be added during the course):

- SpiderLing web crawler for linguistics: http://nlp.fi.muni.cz/trac/spiderling .
- BootCat web crawler for corpus construction: http://bootcat.sslmit.unibo.it/ .
- justext boilerplate removal tool: https://code.google.com/p/justext/ .
- onion deduplication tool for text corpora: http://code.google.com/p/onion/ .
- multiple tools from the Stanford Natural Language Processing Group: http://nlp.stanford.edu/software/index.shtml .

- FreeLing tools suite: http://nlp.lsi.upc.edu/freeling/ .
- multiple tools from the University Centre for Computer Corpus Research on Language (Lancaster University): http://ucrel.lancs.ac.uk/tools.html .
- TreeTagger: http://www.cis.uni---muenchen.de/~schmid/tools/TreeTagger/ .
- Shalmaneser: A Shallow Semantic Parser: http://www.coli.uni---saarland.de/projects/salsa/shal/ .
- Praat and ELAN for audio and video annotation: http://www.fon.hum.uva.nl/praat/ and https://tla.mpi.nl/tools/tla---tools/elan/ .
- EXMARaLDA tools for oral corpora: http://www.exmaralda.org/en/ .
- Praaline system for speech corpora: https://github.com/porcupinebrux/Praaline .
- Ant series tools: http://www.laurenceanthony.net/software.html .
- Python and the Natural Language Toolkit (NLTK) package: https://www.python.org/ and http://www.nltk.org/ .
- IMS Open Corpus Workbench and CQPWeb: http://cwb.sourceforge.net/ .
- R and R---Studio: http://www.r---project.org/ and http://www.rstudio.com/ .
- Additional corpus tools and resources as required by specific project work.

Description of lab activities and assignments: Students will be given a task/experiment to perform in each weekly lab session and will write a lab report on their experiences with the project from that session. In these lab reports the students will assess their work with the tools, including: a description of the tools and methods being investigated that week, a summary of the task/experiment performed, a discussion of the outcomes achieved, and an analysis of the successes and failures of the procedures followed. The exact nature of each assignment will vary depending on the specific tools and methods in use and the type of task/experiment to be performed with those tools and methods. Some of the lab work will be individual assignments, some will be group assignments. Full details for the assignments will be given each week. Due dates for the assignments are listed in the course schedule.

Undergraduate students will write a lab report for each of the 11 ***topics*** covered, for a total of 11 reports over the semester. For topics that last two weeks, undergraduate students will choose one lab session from the two weeks to report on.

Graduate students will write a lab report for each of the 15 ***lab sessions*** for a total of 15 reports over the semester.

There will be no examinations in this lab and no final examination.

Course evaluation and grading: Course grades will be calculated as follows:

The lab reports (15 for graduate students, 11 for undergraduates) represent 100% of the graded material in this course. Each lab report will be equally weighted, such that each is worth an identical portion of the final grade. Undergraduates will receive a midterm grade in week 7 of the course.

Grad student grading scale:
100---90% = A ; 89---80% = B ; 79---70% = C ; 69% and below = E

Undergraduate grading scale:
100---90% = A ; 89---80% = B ; 79---70% = C ; 69---60% = D ; 59% and below = E

Course policies:

Submission of assignments: Students will submit all written lab reports in digital form through the course website. Since assignment submission is digital, there are no exceptions to the due dates listed in the course schedule, unless an excused absence prevents completion and submission by the scheduled due date. Any lab report that is submitted after the due date without an excused absence will be assigned a grade of "0" (zero).

Attendance policy: Attendance will not be a separately graded element in this course, but your attendance habits will likely affect your grade because absence from the lab means that you miss the opportunity to contribute to and learn from the collaborative lab experience. If you miss a lab session for any reason, it is professional courtesy to let me know the general circumstances of your absence, and it is your responsibility to find out what was covered in that session and ensure that you understand the information and concepts discussed in order to be able to complete the lab work on your own. Get notes from your classmates or come to see me to find out what you missed. In addition, I expect everyone to come to the lab on time and to stay for the full duration of the lab session. Again, any missed portion of a lab period is a missed chance to better understand and assimilate the material.

Excused absences: Students need to notify the professor of absences prior to class when possible. S.R. 5.2.4.2 defines the following as acceptable reasons for excused absences: (a) serious illness, (b) illness or death of family member, (c) University---related trips, (d) major religious holidays, and (e) other circumstances found to fit "reasonable cause for nonattendance" by the professor.

Students anticipating an absence for a major religious holiday are responsible for notifying the instructor in writing of anticipated absences due to their observance of such holidays no later than the last day in the semester to add a class. Information regarding dates of major religious holidays may be obtained through the religious liaison, Mr. Jake Karnes (859---257---2754).

Students are expected to withdraw from the class if more than 20% of the classes scheduled for the semester are missed (excused or unexcused) per university policy.

Verification of absences: Students may be asked to verify their absences in order for them to be considered excused. Senate Rule 5.2.4.2 states that faculty have the right to request "appropriate verification" when students claim an excused absence because of illness or death in the family. Appropriate notification of absences due to university---related trips is required prior to the absence.

Academic integrity: Per university policy, students shall not plagiarize, cheat, or falsify or misuse academic records. Students are expected to adhere to University policy on cheating and plagiarism in all courses. The minimum penalty for a first offense is a zero on the assignment on which the offense occurred. If the offense is considered severe or the student has other academic offenses on their record, more serious penalties, up to suspension from the university may be imposed.

Plagiarism and cheating are serious breaches of academic conduct. Each student is advised to become familiar with the various forms of academic dishonesty as explained in the Code of Student Rights and Responsibilities. Complete information can be found at the following website: http://www.uky.edu/Ombud. A plea of ignorance is not acceptable as a defense against the charge of academic dishonesty. It is important that you review this information as all ideas borrowed from others need to be properly credited.

Part II of *Student Rights and Responsibilities* (available online at http://www.uky.edu/StudentAffairs/Code/part2.html) states that all academic work, written or otherwise, submitted by students to their instructors or other academic supervisors, is expected to be the result of their own thought, research, or self---expression. In cases where students feel unsure about the question of plagiarism involving their own work, they are obliged to consult their instructors on the matter before submission.

When students submit work purporting to be their own, but which in any way borrows ideas, organization, wording or anything else from another source without appropriate acknowledgement of the fact, the students are guilty of plagiarism. Plagiarism includes reproducing someone else's work, whether it be a published article, chapter of a book, a paper from a friend or some file, or something similar to this. Plagiarism also includes the practice of employing or allowing another person to alter or revise the work which a student submits as his/her own, whoever that other person may be.

Students may discuss assignments among themselves or with an instructor or tutor, but when the actual work is done, it must be done by the student, and the student alone. When a student's assignment involves research in outside sources of information, the student must carefully acknowledge exactly what, where and how he/she employed them. If the words of someone else are used, the student must put quotation marks around the passage in question and add an appropriate indication of its origin. Making simple changes while leaving the organization, content and phraseology intact is plagiaristic. However, nothing in these Rules shall apply to those ideas which are so generally and freely circulated as to be a part of the public domain (Section 6.3.1).

***Please note**:* Any assignment you turn in may be submitted to an electronic database to check for plagiarism.

***Accommodations due to disability*:** Any student with a disability who is taking this course and needs classroom or exam accommodations should contact the Disability Resource Center, 257-2754, Suite 407, Multidisciplinary Science Building, 725 Rose Street, dtbeac1@uky.edu.

Schedule: The following dates are approximate and are subject to change based on our work with the material.

Week	***Topic***	***Task/Experiment***
Week 1	Corpus construction.	Data collection from print.
Week 2	Corpus construction.	• Week 1 report due (*everyone*). Data collection from audio/video.
Week 3	Corpus construction.	• Week 2 report due (*everyone*). Data collection from the web.
Week 4	Corpus construction.	• Week 3 report due (*everyone*). Data annotation and mark---up – written corpora #1.
Week 5	Corpus construction.	• Week 4 report due (*grad students only*). Data annotation and mark---up – written corpora #2.
Week 6	Corpus construction.	• Week 5 report due (*grad students*). • Week 4---5 report due (*undergraduates*). Data annotation and mark---up – oral corpora #1.
Week 7	Corpus construction.	• Week 6 report due (*grad students only*). Data annotation and mark---up – oral corpora #2.
Week 8	Corpus investigation.	• Week 7 report due (*grad students*). • Week 6---7 report due (*undergraduates*). Frequencies, keywords, concordances.
Week 9	Corpus investigation.	• Week 8 report due (*everyone*). Collocations, colligations, n---grams.
	Spring Break	*no class*
Week 10	Corpus investigation.	• Week 9 report due (*everyone*). Semantics, pragmatics, sentiment analysis.
Week 11	Corpus investigation.	• Week 10 report due (*everyone*). Data---driven pattern matching.
Week 12	Corpus data analysis.	• Week 11 report due (*everyone*). Statistical models and measures #1.
Week 13	Corpus data analysis.	• Week 12 report due (*grad students only*). Statistical models and measures #2.
Week 14	Corpus data analysis.	• Week 13 report due (*grad students*). • Week 12---13 report due (*undergraduates*). Visualization for data analysis #1.
Week 15	Corpus data analysis.	• Week 14 report due (*grad students only*). Visualization for data analysis #2.
Finals week		• Week 15 report due (*grad students*). • Week 14---15 report due (*undergraduates*). *No final exam.*