

|   |  |
|---|--|
| <p><b>Fall 2024</b></p> <p><b><u>LIN4761C/LIN6765/SSA4930 Methods in language documentation</u></b></p> <p><b>Time:</b> MWF 7<sup>th</sup></p> <p><b>Place:</b></p> | <p><b>Instructors</b></p> <p>James Essegbey</p> <p>Office Hours: Mondays and Wednesdays<br/>10:00-11:00 and by appointment<br/>Office: 342 Pugh Hall<br/>Phone: 352-846-2431<br/>Email: <a href="mailto:essegbey@ufl.edu">essegbey@ufl.edu</a></p> <p>Sarah Moeller</p> <p>Office Hours: MW 10:30-11:30<br/>Office: 4017 Turlington Hall<br/>Phone: 352-294-7449<br/>Email: <a href="mailto:smoeller@ufl.edu">smoeller@ufl.edu</a></p> |
|---|--|

### **Course objectives**

David Nathan writes “language documentation, as a new field operating in a largely digital environment, can aspire to exploit all the capabilities of new technologies.” The aim of this course is to equip students with skills to exploit these technologies. Students will learn about choosing appropriate audio and video recording tools and how to get the best results from the ones they choose. They will also learn to process video and audio recording using EUDICO Linguistic Annotator (ELAN) and FieldWork Language Explorer (FLEx). Students will learn about metadata and preparing tests for archiving. They will learn about choosing from existing (nonlinguistic) stimuli and creating supplementary ones to address their specific research needs, and using computational methods to process data.

**Course Material:** The materials for the course are class notes, articles in journals, book chapters. Some of the articles are available for download at your course site in CANVAS while others are posted at ARES.

### **Requirements**

Final grading is based on an assignment, a test, and a project. While the 4000-level students will do the project work in pairs, 6000-level students will do it alone. The breakdown is as follows:

|   |        |
|---|--------|
| Assignment (A-E = 5 pts; F = 10 pts)                                | 35 pts |
| Group (4000-level) / Individual (6000-level) Project & Presentation | 50 pts |
| Reflection paper  | 15 pts |

### **Week 1. Syllabus (Dr. Essegbey & Dr. Moeller)**

**Fri, Aug 23.** Review of Syllabus

### **Week 2. Introduction (Dr. Essegbey)**

**Mon, Aug 26.** What is language documentation?

**Reading:** Current issues in language documentation by Peter Austin (1)

**Wed, Aug 28.** Language documentation (continued)

**Reading:** Language documentation 20 years on by Peter Austin (2)

**Fri, Aug 30.** Documentation and revitalization

**Reading:** Language documentation and language revitalization—some methodological consideration. By Peter A. Austin and Julia Sallabank (3)

**Week 3. Ethics (Dr. Moeller)**

**Mon, Sept 2. Labor Day. No classes**

**Wed, Sept 4.** Ethical research.

**Reading:** Bowern, Claire. 2008. Chapter: Ethical field research (4)

**Fri, Sept 6.** Ethics in AI

**Reading:** Marie-Odile Junker. 2024. [Data-mining and Extraction: the gold rush of AI on Indigenous Languages](#). In *Proceedings of the Seventh Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 52–57, St. Julians, Malta. Association for Computational Linguistics. (5)

**Week 4. Audio (Dr. Essegbey and Dr. Moeller)**

**Mon, Sept 9. Audio – Microphones**

**Reading:** Sound and unsound practices in documentary linguistics: towards an epistemology for audio by David Nathan (6)

Sound recording: microphones by David Nathan (7)

**Wed, Sept 11.** Audio techniques

**Reading:** Audio and video recording techniques for linguistic research, by Anna Margetts and Andrew Margetts (8)

**Fri, Sept 13.** Processing audio: Audacity

**Reading :** <https://userguiding.com/blog/audacity-tutorial/> (9)

**Week 5. Audio & Video (Dr. Essegbey and Dr. Moeller)**

**Mon, Sept 16.** Practice audio recording and trimming audio data

**Wed, Sept 18.** Review audio data

**Assignment A due.**

**Fri, Sept 20.** Video

**Reading:** Audio and video recording techniques for linguistic research, by Anna Margetts and Andrew Margetts (8)

**Reading:** The role of digital video in language documentation by Louise Ashmore (10)

**Week 6. Video (Dr. Essegbey)**

**Mon, Sept 23.** Gesture

**Reading:** Reasons for documenting gestures and suggestions for how to go about it. By Mandana Seyfeddinipur (11)

**Wed, Sept 25.** Practice video recording

**Assignment B due (group/pairs).**

**Fri, Sept 27.** Review video recording

**Week 7. Annotation (Dr. Essegbey and Dr. Moeller)**

**Mon, Sept 30.** Annotation

**Reading:** Linguistic annotation by Eva Schultze-Berndt (12)

**Wed, Oct 2.** Orthography

**Reading:** Is this my language? Developing a writing system for an endangered-language community by James Essegbey (13)

**Fri, Oct 4.** AI and data annotation

**Reading:** Geoffrey Leech. 2004. "Developing Linguistic Corpora: A Guide to Good Practice" Chapter 1 [Adding Linguistic Annotation](#) (14)

**Week 8. ELAN and video/audio transcription (Dr. Essegbey)**

**Mon, Oct 7.** ELAN

**Download:** <https://tla.mpi.nl/tools/tla-tools/elan/download>

**Wed, Oct 9.** Transcription Mode

**Reading:** <http://ideophone.org/transcription-mode-in-elan> (15)

**Fri, Oct 11.** Practice transcription in ELAN

**Assignment C due.**

**Week 9. Metadata (Dr. Essegbey & Dr. Moeller)**

**Mon, Oct 14.** Different types of metadata

**Reading:** Reconceiving metadata: language documentation through thick and thin. By David Nathan and Peter Austin (16)

**Wed, Oct 16.** Lameta and/or SayMore

**LaMeta Download & Reading:**

<https://sites.google.com/site/metadatatooldiscussion/home?pli=1> (17)

**SayMore Download:** <https://software.sil.org/saymore/download/>

**SayMore Reading:**

<https://drive.google.com/drive/folders/0B2TFy02zqdRwOE9sdEZaNF1pcHM?resourcekey=0-cJen8e6Fil3GKMiJwYSCzw&usp=sharing> (18)

**Fri, Oct 18. Homecoming. No classes**

**Week 10. Computational methods with experiment (Dr. Moeller)**

**Mon, Oct 21.** Introduction to NLP

**Assignment D due.**

**Reading:** Luke Gessler and Katharina von der Wense. 2024. [NLP for Language Documentation: Two Reasons for the Gap between Theory and Practice](#). In *Proceedings of the 4th Workshop on Natural Language Processing for*

*Indigenous Languages of the Americas (AmericasNLP 2024)*, pages 1–6, Mexico City, Mexico. Association for Computational Linguistics. (19)

**Wed, Oct 23.** Training NLP machine learning model with documentary data

**Reading:**

[https://www.voanews.com/a/usa\\_google-explains-why-app-cant-translate-most-native-american-languages/6204275.html](https://www.voanews.com/a/usa_google-explains-why-app-cant-translate-most-native-american-languages/6204275.html) (20)

**Fri, Oct 25.** Practice preparing/cleaning data for NLP

**Week 11. Elicitation Methods (Dr. Essegbey)**

**Mon, Oct 28.** Picture Elicitation

**Reading:** Frog story

**Wed, Oct 30.** Video elicitation

**Reading:** Causality Video

**Fri, Nov 1.** Gesture elicitation

**Reading:** Kita, Sotaro and Essegbey, James. 2001. Pointing left in Ghana: how a taboo on the use of the left hand influences gestural practices. *Gesture* 1 (1), 73-95 (21)

**Week 12. Working with FLEx (Dr. Moeller)**

**Mon, Nov 4.** Introduction to FLEx

**Assignment E due.**

**Download:** <https://software.sil.org/fieldworks/>

**Wed, Nov 6.** Configuring Font, Orthographies, and Interlinear Gloss Texts in FLEx

**Fri, Nov 8.** Practice configuring Text & Words in FLEx

**Week 13. Computational methods 2 (Dr. Moeller)**

*Mon, Nov 11. Veteran's Day, No classes*

**Wed, Nov 13.** Human/Machine-in-the-Loop

**Assignment F due.**

**Reading:** Forbes et al. 2022. "[Dim Wihl Gat Tun: The Case for Linguistic Expertise in NLP for Underdocumented Languages](#)" (22)

**Fri, Nov 15.** Review results of model improvements

**Week 14. Work on Project and Presentation (No Classes)**

**Week 15. Thanksgiving Break**

**Week 16. Presentations of Project**

**Mon, Dec 2**

**Wed, Dec 4**

**Grading Scale**

93-100 (A); 90-92 (A-); 87-89 (B+); 83-86 B; 80-82 (B-); 77-79 (C+); 73-76 (C); 70-72 (C-); 67-69 (D+); 63-66 (D); 60-62 (D-); 59 or below (E)

A grade of Incomplete (I) will not be issued under any circumstance.

### **Grading Policy**

Information on current UF grading policy can be found at:

<https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx>

### **Grading Scale**

93-100 (A); 90-92 (A-); 87-89 (B+); 83-86 B; 80-82 (B-); 77-79 (C+); 73-76 (C); 70-72 (C-); 67-69 (D+); 63-66 (D); 60-62 (D-); 59 or below (E)

A grade of Incomplete (I) will not be issued under any circumstance.

### **Grading Policy**

Information on current UF grading policy can be found at:

<https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx>

### **Students with disabilities**

The University of Florida provides high-quality services to students with disabilities, and you are encouraged to take advantage of them. Students with disabilities needing academic accommodations should 1) Register with and provide documentation to Disability Resources (352-392-8565), and 2) Bring a letter to the instructor from Disability Resources indicating that you need academic accommodations. Please do this as soon as possible, preferably within the first week of class.

### **Class Attendance and Make-Up Policy**

Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found in the online catalog at:

<https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx>

Class attendance is expected. Excused absences are consistent with university policies in the undergraduate catalog and require appropriate documentation.

Students who can demonstrate that they were unable to submit an assignment by the deadline due to an excused absence and who can provide appropriate documentation for the absence will be given a reasonable period of time to make up the late work.

Covid-19 statements recommended by UF: If you are not vaccinated, get vaccinated. Vaccines are readily available and have been demonstrated to be safe and effective against the COVID-19 virus. Visit [one.ufl](http://one.ufl) for screening/testing and vaccination opportunities.

If you are sick, stay home. Please call your primary care provider if you are ill and need immediate care or the UF Student Health Care Center at 352-392-1161 to be evaluated.

Course materials will be provided to you with an excused absence, and you will be given a reasonable amount of time to make up work.

## **Students Requiring Accommodations**

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565 <https://disability.ufl.edu/get-started/>) by providing appropriate documentation. Once registered, students will receive an accommodation letter that must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

## **Course Evaluation**

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <https://gatorevals.aa.ufl.edu/students/>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <https://ufl.bluera.com/ufl/>. Summaries of course evaluation results are available to students at <https://gatorevals.aa.ufl.edu/public-results/>.

## **Class Demeanor**

Students are expected to arrive to class on time and behave in a manner that is respectful to the instructor and to fellow students. Please avoid the use of cell phones and restrict eating to outside of the classroom. Opinions held by other students should be respected in discussion, and conversations that do not contribute to the discussion should be held at minimum, if at all. Laptops can be used for note taking.

## **Materials and Supplies Fees**

There are no additional fees for this course.

## **University Honesty Policy**

UF students are bound by The Honor Pledge that states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (<https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor in this class.

## **Counseling and Wellness Center**

Contact information for the Counseling and Wellness Center: <http://www.counseling.ufl.edu/cwc/Default.aspx>, 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

### **U Matter, We Care:**

If you or someone you know is in distress, please contact [umatter@ufl.edu](mailto:umatter@ufl.edu), 352-392-1575, or visit U Matter, We Care website (<https://umatter.ufl.edu/>) to refer or report a concern and a team member will reach out to the student in distress.

### **Writing Studio**

The writing studio is committed to helping University of Florida students meet their academic and professional goals by becoming better writers. Visit the writing studio online at <http://writing.ufl.edu/writing-studio/> or in 302 Tigert Hall for one-on-one consultations and workshops.

### **Regulation around In-Class Recording:**

Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited. Specifically, students may not publish recorded lectures without the written consent of the instructor.

A “class lecture” is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a guest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session.

Publication without permission of the instructor is prohibited. To “publish” means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code.