

IDS 2935: Language and Computers

Quest 2

I. General Information

Class Information

- Spring 2023
- **Lecture Time:** MW 9:35am
- **Lecture Classroom:** Turlington 005
- **Discussions:** Fridays; various times and locations

Instructor

- Sarah Moeller
- **Office:** Turlington 4017
- **Office Hours:** M 10:40-11:45; W 3-3:50
- smoeller@ufl.edu

Teaching Assistants

- Alexis Davis
- Office: Turlington Hall
- [Office hours]
- davisa2@ufl.edu
- Shengyu Liao
- **Office:** Dauer Hall 36
- **Office Hours:** T 3:00-4:55 p.m.
- shengyu.liao@ufl.edu

Course Description

Language technology has a profound influence on the way ordinary people use language. This morning, because you speak English, you may have already used voice recognition or predictive text. This course explains what language technology is and how it is available for 1% of the world's languages.

Along the way, we will attempt to answer a big pressing question for our society: Can artificial intelligence ever be inclusive of all 7000+ human languages? We will identify, describe, and explain the cross-disciplinary dimensions of this question which lie partly in social science principles of human communication (e.g. conversational turn-taking), partly in linguistics theory (e.g. how languages build words), and partly in the history of computer science (e.g. ASCII vs. Unicode). Topics include spellcheckers, translation, chatbots, and language learning aids. Topics are explored in the context of globalization, language endangerment, and the recent rapid rise of artificial intelligence.

This course assumes no background in linguistics or computer science. It introduces the fundamentals of language science so that students can peek under the hood of AI. Students will critically analyze digital systems that "understand" human language and related policies and practices. They will explore how communities around the world experience language technology and how it mediates our communication in an increasingly connected world. Homework assignments require students to connect analytical, linguistic, and technological skills to the human and AI behavior they encounter in their personal and professional life. The writing requirement asks them to communicate a problem or challenge that arises when we try teaching computers how to process human language and to develop a reasoned critique or recommendation.

Quest and General Education Credit

- Quest 2
- Social & Behavioral Sciences
- Writing Requirement (WR) 2000 words

This course accomplishes the [Quest](#) and [General Education](#) objectives of the subject areas listed above. A minimum grade of C is required for Quest and General Education credit. Courses intended to satisfy Quest and General Education requirements cannot be taken S-U.

Required Readings and Works

- Dickinson, Markus, Chris Brew, and Detmar Meurers. 2013. *Language and Computers*. Wiley-Blackwell. https://ufl-flvc.primo.exlibrisgroup.com/permalink/01FALSC_UFL/175ga98/alma990368587650306597 (abbreviated as **LaC** on schedule)
(The textbook is available as an e-book through the University of Florida library system that a very limited number of users can either read online or download a subsection up to 51 pages every 24 hours).
- Required writing guide: “The Only Guide to Essay Writing You’ll Ever Need”. Grammarly Blog. Online: <https://www.grammarly.com/blog/essay-writing/>
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.*
- Required citation guidelines:
https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/index.html
- All additional readings will be available on Canvas.
- Materials and Supplies Fees: n/a

II. Graded Work

Description of Graded Work

The grade for this course is based on in-class exercises, 5 homework assignments, 1 essay with a minimum of 2000 word, and 2 exams.

Participation	5%
Midterm exam	15%
Writing assignment	20%
Final exam	20%
Homework/Self-reflection	40%

Participation: Homework and the exams address the material covered in class as well as readings, so good attendance is essential. Group and individual exercises are assigned from time to time in class and these exercises count for a sizable portion of the participation grade. The grade for an in-class exercise

may be excused only if attendance has been excused by the instructor. Inform the instructor *beforehand* if you cannot make a class time.

Homework Assignments: Homeworks will be assigned about every two weeks. Students are encouraged to discuss assignments with classmates and reason out problems together, but all submitted work must be their own! Late assignments are deducted 10% of the assessed grade for every day after the Canvas deadline (not counting Sundays or your preferred day for religious observance / work-life balance).

Midterm Exam: The mid-term exam will review material covered during Weeks 1-7.

Final Exam: The final exam (see one.uf.edu for date) is cumulative, but with a greater emphasis on the content covered after the midterm.

Writing Assignment (satisfies E2 writing requirement): Students will submit a 2250-word (+/- 10%) essay that critically analyzes an issue or challenge related to some dimension of language technology's impact on human society and human languages. The essay must propose or critique an approach, policy, or practice that attempts to address their chosen issue or challenge. To do this, the essay must satisfy the points below, with a minimum of one paragraph per point. Each paragraph must have a clear thesis statement that is supported by specific examples and/or illustrations. Students are encouraged to employ charts and figures with descriptive captions and to make educated use of quotes, citations, and create a bibliography. They must use the APA (7th ed.) citation style. Also, don't forget a good title and the introductory and conclusion paragraphs! Feedback on writing skills will be provided by the end of finals.

- **What is a current issue or challenge related to whether computers can store, represent, or process ("understand") all human languages or all ways of using a given language?** This could be a problem or question specific to a task or a type of technology or it may be a challenge or trend found in multiple related areas. Possible current issues or challenges may be found in the assigned readings or in newspapers, blogs, social media, etc. (Be sure to cite!), or something you yourself have experienced.
Example thesis sentences: *"Computer-assisted language learning technology are rarely available for languages spoken by less than X number of speakers."* *"My own experiments (which I will describe and illustrate) have shown that Y machine translation system performs better between A and B related languages than between C and D unrelated languages."* *"Z researchers say that the quality of language technology is dependent on having very large amounts of data."*
- **In what concrete ways has this issue or challenge been addressed or solved?** Be as specific as possible. You may find responses or solutions in headlines, social media, or specialized publications. The issue or challenge may be addressed in the form of events, laws, new products, public discussions in published or social media, sponsored research, etc. Try to answer questions such as (but not limited to): How are people trying to solve or exploit the issue or challenge? Do proposed solutions exist? Do they seem possible with current resources and capabilities of language technology? You may also address related questions: Who do these responses come from? Why? What issues and challenges do proposed solutions or concrete methods bring with them? If you had to predict, where do related trends seem to be headed?
Examples thesis sentences: *"The language learning company A claims its products are teaching and revitalizing endangered languages."* *"W team of researchers say they improved machine translation with X number of new techniques but these techniques require Y and Z."* *"A, B, and C*

companies use creative, non-monetary ways to compensate users of their “free” systems for the language data that those users are unknowingly providing to train intelligent systems.”

- **How does the issue or challenge, and society’s responses to it, impact society? Are the responses and proposed solutions to the issue or challenge good or bad for individuals or society as a whole?** Here, you must include other’s thoughts and opinions. You should examine information and opinions expressed in popular media, newspaper headlines, researchers’ blogs, etc. Or you may conduct interviews with friends and family (include your interview questions in an appendix). Be sure to give proper citation and attribution.
Example thesis sentence: *“Language technology’s dependence on data may threaten personal privacy.”*
- **Self-reflection: How has this issue or challenge impacted you or someone you know? What steps would you personally take to respond to this issue or challenge, both now and in the future? How do you think your feeling or response to this issue or challenge changed since taking this course?** Express your opinion and outline recommendations. Reflect on how the key themes and principles, methodologies, and case studies covered in this course intersect with this issue and your own life and future career.

Grading Scale

For information on how UF assigns grade points, visit: <https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/>

A	94 – 100%		C	74 – 76%
A-	90 – 93%		C-	70 – 73%
B+	87 – 89%		D+	67 – 69%
B	84 – 86%		D	64 – 66%
B-	80 – 83%		D-	60 – 63%
C+	77 – 79%		E	<60

Grading Rubric(s)

Writing Assessment Rubric

	SATISFACTORY (Y)	UNSATISFACTORY (N)
CONTENT	Exhibits some evidence of ideas that respond to the topic with complexity. Critically evaluates and synthesizes sources and provides an adequate discussion with basic understanding of sources.	Either includes a central idea(s) that is unclear or off-topic or provides minimal or inadequate discussion of ideas. May lack sufficient or appropriate cited sources.
ORGANIZATION AND COHERENCE	Exhibits identifiable structure, including a clear central thesis statement and thesis statement for each paragraph. At the weak end of Satisfactory, it may require readers to work to follow progression of ideas.	Paragraphs lack clearly identifiable organization. May lack coherent sense of logic for associating and organizing ideas. May lack transitions to guide reader.
ARGUMENT AND SUPPORT	Uses persuasive and confident presentation of ideas, strongly supported with evidence. Examples illustrate thesis statements. At the weak end of Satisfactory, may provide only generalized discussion of ideas or provides adequate discussion but relies on weak support or poorly chosen examples.	Paper makes only weak generalizations, providing little or no support, as in summaries or illustrations without critical analysis or clear connection to central ideas.
STYLE*	Uses a writing style and word choice appropriate to the context, genre, and discipline. Sentences display complexity and logical sentence structure. On the weak end, documents display a less precise use of vocabulary and an uneven use of sentence structure or use a writing style that occasionally veers in word choice or tone appropriate to the context, genre, and discipline.	Author relies on writing style and word choice that is inappropriate for the context, genre, or discipline. Sentences may be overly long or short or have awkward construction. Author may use words incorrectly.
MECHANICS*	Features correct and error-free presentation of ideas. Contains no errors caught by a typical spellchecker. At the weak end of the Satisfactory, may contain word choice, punctuation, or grammatical errors that do not muddy the paper's argument or points.	Contains so many mechanical or grammatical errors that they impede the reader's understanding or severely undermine the writer's credibility.

- The Writing Requirement (WR) ensures students both maintain their fluency in writing and use writing as a tool to facilitate learning.
- The instructors will evaluate before the end of the course all written assignments with respect to grammar, punctuation, clarity, coherence, and organization and provide feedback.
- WR course grades have two components. To receive writing requirement credit, a student must receive a grade of C or higher on the course and a satisfactory completion of the writing component of the course.

Participation & Self-Reflection Rubric

	High Quality	Average	Needs Improvement
Informed	Shows evidence of having done the assigned work		Does not show evidence of having done the assigned work
Thoughtful	Shows evidence of having understood and considered issues raised		Does not show evidence of having engaged with or considered issues raised
Considerate	Takes perspective of others into account		Does not take perspective of others into account

III. Annotated Weekly Schedule

Week	Topics, Homework, and Assignments	Required Readings	Assignments (Due dates on Canvas)
Week 1	Topic: Can Computers “Read” All Alphabets? Summary: How many ways do humans write? Why did computers abandon the ASCII table for Unicode?	LaC pp. 1-17	
Week 2	Topic: Can Computers “Hear” All Languages? Summary: Why does(n’t) your phone understand you? How do computers connect language with speech-to-text (ASR) and text-to-speech (TTS) systems?	LaC pp. 17-29 (including Under the Hood 2)	
Week 3	Topic: Can Computers Spell Any Language? Summary: How many kinds of spelling errors do humans make? How are spellcheckers built and what/who do they represent?	LaC pp. 33-44	DUE: Homework 1
Week 4	Topic: Can Computers Help Anyone Write? Summary: What unique language structures do writing assistants need to know? How do computers distinguish syntactic (structural) and semantic (meaning) errors?	LaC pp. 49-65	Final paper discussed
Week 5	Topic: Can Computers Teach Us Languages? Summary: Do computer-aided language learning systems give a more effective learning experience? How might they increase economic opportunities while creating “killer” languages?	LaC pp. 69-87	
Week 6	Topic: What Languages Use the Internet? Summary: What are the tools and techniques of search engines? How does search technology deal with the differences between structured, unstructured, and semi-structured data?	LaC pp. 91-120	DUE: Homework 2
Week 7	Topic: Midterm review & Midterm	LaC pp. 91-120	
Week 8	Topic: How do Computers Learn? What is AI? Summary: What are the theoretical and formal fundamentals of AI for natural language processing? How is the history of AI tied to linguistics? Why can computers “understand” only a small percentage of the world’s languages?	LaC pp. 127-133	
Week 9	Topic: Can Computers Understand Emotion?	LaC pp. 133-151	

Week	Topics, Homework, and Assignments	Required Readings	Assignments (Due dates on Canvas)
	Summary: How do computers recognize the author of a document and the author's attitude (sentiment)? How does this help fill your spam folder?		
BREAK			
Week 10	Topic: How We Talk Summary: How do human conversations happen? What are Grice's Maxims and why do we flout them when we talk?	LaC pp. 153-166	DUE: Homework 3 (Paper Draft)
Week 11	Topic: Who Do Computers Like to Talk With? Summary: How do computers "converse"? How are chatbots built?	LaC pp. 166-177	DUE: Homework 4
Week 12	Topic: How Many Languages Do We Speak? Summary: What ways do languages differ in structure? Why is machine translation important but not always appropriate? How does society influence translation?	LaC pp. 181-198	
Week 13	Topic: Are Computers Multilingual? Summary: Why was machine translation the first task attempted by AI? What makes translation hard?	LaC pp. 198-209	DUE: Homework 5
Week 14	Topic: Computers and Languages that Die Summary: Why are so many human languages in danger of disappearing? Why is language technology "biased" against minority languages? Could technology be leveraged to revitalize endangered and extinct languages?	Reading on Canvas	
Week 15	Topic: Impact of Language Technology. Summary: Does the nature and quality of communication change with language technology? What aspects of language may be beyond a computer's grasp?	LaC pp. 215-219	DUE: Writing Requirement Paper

IV. Student Learning Outcomes (SLOs)

At the end of this course, students will be expected to have achieved the [Quest](#) and [General Education](#) learning outcomes as follows:

- **Content:** Students demonstrate competence in the *terminology, concepts, theories and methodologies*. They identify, describe, and explain key themes, principles, and terminology; the

history, theory and/or methodologies used; and social institutions, structures and processes within the discipline(s).

- Identify, describe and explain key theoretical terms and concepts that apply to languages around the world and the communities that speak them: syntax, morphology, pragmatics, history of encoding written language, language endangerment and conservation, principles and politics of translation (**S**)
- Identify and understand key structures and terms in computer science sufficiently well to explain how computers are able to store, represent, and process written and spoken data, e.g. graphs and trees, formal grammars, language classification, machine learning algorithms (**S, Quest 2**)
- Identify and describe endangered languages around the world and in the U.S.; explain how languages become endangered and why minority language communities remain under-served by technology (**Quest 2**)

Assessments: Exams, exercises, homeworks

- **Critical Thinking:** *Students carefully and logically analyze information from multiple perspectives and develop reasoned solutions to problems within the discipline(s). They apply formal and informal qualitative or quantitative analysis effectively to examine the processes and means by which individuals make personal and group decisions. Assess and analyze ethical perspectives in individual and societal decisions.*

- Analyze the theory and practice of human language technology, including the scientific, commercial, and social goals for which it can be applied (**S, Quest 2**)
- Critically examine and evaluate how the limitations and ethical challenges of the language technology may impact a speech community's educational, economic, and political involvement (**S, Quest 2**)
- Evaluate the capabilities of a specific language technology system for processing a specific language or language variety, linguistic structure, or writing system used by a community around the world or in the U.S. (**Quest 2**)

Assessments: exercises, homeworks

- **Communication:** *Students communicate knowledge, ideas and reasoning clearly and effectively in written and oral forms appropriate to the discipline(s).*

- Identify and explain in writing at least one current issue at the intersection of technology, language, and social science using concepts, terms, and ideas covered in this course (**S, Quest 2**)
- Develop and analytically support one's own conclusions, and critically assess others' opinions, about the design, capabilities and behavior of human language and the technology that stores, represents, and processes languages (**S, Quest 2**)

Assessments: writing requirement, homeworks, exercises

- **Connection:** *Students connect course content with meaningful critical reflection on their intellectual, personal, and professional development at UF and beyond.*

- Connect the structures and concepts that are "under the hood" in language technology to other technology regularly encountered in adult and professional life (**Quest 2**)
- Reflect on social and ethical considerations related to language technologies and connect them to broader societal issues, such as privacy, equity, and the nature of artificial intelligence (**Quest 2**)

- Compare ways that technology can increase the threat to and also help maintain languages spoken by under-resourced communities around the world and in the U.S. (Quest 2)

Assessments: exercises, homeworks, writing requirement

V. Quest Learning Experiences

1. Details of Experiential Learning Component

Experiential learning connects the concepts, history, themes, and terminology covered by lectures and readings to concrete applications, or to related issues or skills that students encounter regularly in our lives. Most exercises in this course are experiential learning activities. For example, after learning how computers encoded spoken language and what limitations computer have when processing spoken language, students will be challenged to “confuse” a voice recognition system such as Amazon Alexa. They will observe, report, and discuss the results.

Some homework assignments guide students through an experiential learning component. The first part contains one or more activities that will demonstrate the student’s understanding of the course material. Typically, one activity requires hands-on experiential learning with commercial language technology. For example, after reviewing basic grammatical terms and concepts, and being introduced to word-building strategies in various languages, students may perform a basic linguistic analysis of an unfamiliar language, and based on that analysis, translate a sentence from that language into English. Or, after learning how computers “hear” language, students may try out Amazon or Google’s automatic speech recognition systems.

2. Details of Self-Reflection Component

Homeworks with self-reflection have two parts. The first is a learning activity. The second part guides students to reflect on their learning and to contextualize and deepen their connection to the course content. For example, after analyzing their attempts to “break” an automatic speech recognition system, students will reflect on their experience and how it might relate to the everyday experience of non-native or “non-standard” speakers encountering technology trained to “hear” one way of speaking. Or, after learning about multilingualism around the world and about machine translation systems, students may be given an exercise in class to translate foreign words with a dictionary. Then students will reflect in writing or group discussions about aspects that make the task difficult, what aspects might be difficult for a computer, and how such difficulties might affect people who would like to be involved in their government in countries where the official language is different than their native tongue.

The writing requirement also provides an opportunity to reflect on and communicate conclusions about the design, capabilities and behavior of human languages and the technology that processes them. Students will consider content from readings and lectures and explore media to identify a current issue at the intersection of language, technology, and social realities. Students must evaluate, reflect on, and communicate social and ethical considerations related to this issue.

VI. Required Policies

Attendance 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 at:

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

Students Requiring Accommodation

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting

<https://disability.ufl.edu/students/get-started/>. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible.

UF Evaluations Process

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/>.

University Honesty Policy

UF students are bound by The Honor Pledge which 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 or TAs in this class.

Religious Observances

A student should inform the instructor if the religious observances of their faith will conflict with class attendance, tests or examinations, or other class activities. This should be done prior to the class or occurrence of that test or activity. We are obligated (and happy) to accommodate that students’ religious observances.

Counseling and Wellness Center

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

Inclusive environment.

Students are encouraged to employ critical thinking and to rely on data and verifiable sources to interrogate all assigned readings and subject matter in this course as a way of determining whether they agree with their classmates and/or their instructor. No lesson is intended to espouse, promote, advance, inculcate, or compel a particular feeling, perception, viewpoint or belief. Students are encouraged to share their viewpoints, data, and sources in class and to speak with the instructor or classmates, in class or privately, about any perceived violation of this policy.

In-Class Recordings

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.