CSCI 680

Affective Computing

William & Mary · Department of Computer Science · Fall 2026 · Graduate research seminar, 3 credits

Meetings	Tuesday & Thursday, 11:00 a.m. – 12:20 p.m.
Room	John E. Boswell Hall 220
Office	Integrated Science Center 4, RM 2381
Term	August 27 – December 3, 2026
Instructor	Ashley (Ye) Gao, Assistant Professor of Computer Science
Email	ygao18@wm.edu
Web	https://lindagaw.github.io/courses/CSCI680-Fall26/index.html
Office hours	Tuesday & Thursday, 12:30 p.m. –2:20 p.m.

1. Course description

Affective computing is the study of systems that recognize, model, generate, and respond to human emotion. This seminar covers the modern, machine-learning-centered version of that field: how emotion is represented in learned features, how it is recognized from speech and other signals under realistic conditions, how it is controlled in generative models, and how the resulting systems fail once they leave the lab.

We read one to two technical papers per meeting, drawn mostly from ICML, NeurIPS, ICLR, ACL/EMNLP, ICASSP, and Interspeech. The first third of the semester builds shared ground — emotion theory and its critiques, corpora and labels, self-supervised speech representations. The rest moves through four active research areas: emotion-aware audio-language models, expressive and controllable speech synthesis, robustness under distribution shift, and interpretability of affective representations. The last week belongs to your projects.

This is a discussion course, not a lecture course. Come having read the paper.

2. Learning outcomes

By December you should be able to:

1. Read a technical paper in affective computing critically — identify its claim, its evidence, its baselines, and the gap between what it demonstrates and what it asserts.
2. Explain the major families of methods for emotion recognition and emotion-controllable generation, and articulate the tradeoffs among them.
3. Present a research paper to a technical audience and lead a substantive discussion of it.
4. Situate a specific research contribution within the broader literature.
5. Design, execute, and report on an original research project, including a defensible experimental protocol.

1. Reason about the validity and ethical stakes of inferring emotional state from behavioral signals.

3. Prerequisites

- Graduate standing in Computer Science, or permission of the instructor.
- Working knowledge of machine learning at the level of CSCI 516 or equivalent — neural network training, optimization, evaluation.
- Comfort with Python and at least one deep learning framework; PyTorch preferred.
- Enough mathematical maturity to follow derivations involving probability, linear algebra, and basic information theory.

No speech background needed. Prior work in speech or signal processing helps, but it isn't required. The audio concepts you need are introduced as they come up — the first time a mel spectrogram or a codec token matters, we stop and explain it.

4. How a meeting runs

The presenter (35–40 min). One student covers the problem, the method, the experiments, and the weaknesses — then facilitates the discussion that follows. You'll present once or twice over the semester; sign-ups happen on day one.

The discussion (~30 min). Everyone else arrives having read the paper. The instructor presents in Weeks 1–2 to model the format and adds context throughout, but the seminar belongs to the students.

The reading response (due 9:00 a.m.). 250–400 words on every paper you aren't presenting: the central claim, one substantive critique, one question for the room. Credit / no-credit. Your two lowest scores are dropped, no questions asked.

5. Grading

Component	Weight
Final project — report and presentation	30%
Paper presentations — one or two over the semester	25%
Reading responses	15%
Seminar participation	15%
Midterm progress report	10%
Project proposal	5%

There is no final exam.

5.1 Reading responses (15%)

Due by **9:00 a.m. on the day of each class you are not presenting**. Roughly 250–400 words, graded credit / no-credit on evident engagement. Each response covers three things:

- A one-paragraph summary of the paper's central claim.
- One substantive critique — a methodological gap, an unconvincing experiment, an unexamined assumption, or a claim the evidence does not support.
- One question you want the seminar to discuss.

Your two lowest scores are dropped, no questions asked. They exist to absorb ordinary life — a conference deadline, a bad week, a flight. Use them without explaining. **Late responses are not accepted**; the whole point is that you read the paper before we discuss it.

5.2 Paper presentations (25%)

You present once or twice over the semester. Sign-ups happen at the first meeting on August 27. Plan for **35–40 minutes** covering the problem, the method, the experiments, and the weaknesses — then facilitate about **30 minutes** of discussion. Facilitating is part of the grade: come with questions the room can't answer from the abstract alone.

The weaknesses section is not optional. Every paper on the schedule was accepted somewhere good, and every one of them oversells something. Finding that thing is most of what this course teaches.

5.3 Final project (30%, plus 15% across proposal and progress report)

An original research project, done individually or in pairs.

Acceptable project types

- A novel method with empirical evaluation.
- A careful reproduction and stress-test of a published result.
- A rigorous empirical study of an open question — for example, how a particular class of model degrades under a specific kind of shift.

Not acceptable: pure literature surveys. If the project doesn't run something and measure something, it isn't a project for this course.

Deliverables

Date	Deliverable	Weight
Fri, Sep 18	Project proposal — 2 pages	5%
Fri, Oct 30	Midterm progress report — 4 pages	10%
Dec 1 & 3	In-class presentations, sessions I and II	part of 30%
Fri, Dec 11	Final report — 8 pages, conference format	part of 30%

Any project collecting new data from human participants needs **IRB review before collection begins**. Build that timeline into your proposal (see §8, Research ethics).

Late project work. Project deliverables may be submitted up to 48 hours late with a 10% penalty. Beyond that, only by prior arrangement. Reading responses are never accepted late.

6. Schedule

Classes begin August 26 and end December 4. Fall Break, Election Day, and Thanksgiving remove three of our meetings; November 24 is held online. *The instructor reserves the right to adjust readings. The calendar dates are fixed.*

Unit I — Foundations: what is being measured?

Emotion theory and its critics, then the corpora everything downstream is trained on.

01 · Thu Aug 27 — Course introduction

Scope of the field, seminar mechanics, how to read and present a technical paper. Presentation sign-ups.

- Picard, *Affective Computing*, MIT Media Lab Perceptual Computing TR 321, 1995 — *skim*

02 · Tue Sep 1 — Can emotion be inferred from behavior?

- Barrett, Adolphs, Marsella, Martinez & Pollak, "Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements," *Psychological Science in the Public Interest* 20(1):1–68, 2019 — *read the executive summary and §1–2 closely, skim the rest*

03 · Thu Sep 3 — Corpora I: acted and elicited emotion

- Busso et al., "IEMOCAP: Interactive Emotional Dyadic Motion Capture Database," *Language Resources and Evaluation* 42(4):335–359, 2008
- Cao et al., "CREMA-D: Crowd-Sourced Emotional Multimodal Actors Dataset," *IEEE TAC* 5(4):377–390, 2014

04 · Tue Sep 8 — Corpora II: naturalistic and in-the-wild data

- Lotfian & Busso, "Building Naturalistic Emotionally Balanced Speech Corpus by Retrieving Emotional Speech From Existing Podcast Recordings," *IEEE TAC* 10(4):471–483, 2019
- Mollahosseini, Hasani & Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," *IEEE TAC* 10(1):18–31, 2019

Unit II — Learned speech representations

What self-supervised models learn from raw audio, and how much of it is affect.

05 · Thu Sep 10 — Self-supervised speech representation learning

- Baevski, Zhou, Mohamed & Auli, "wav2vec 2.0," NeurIPS 2020

06 · Tue Sep 15 — Masked prediction and full-stack pretraining

- Hsu et al., "HuBERT," *IEEE/ACM TASLP* 29:3451–3460, 2021
- Chen et al., "WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing," *IEEE JSTSP* 2022

07 · Thu Sep 17 — What do these representations actually encode?

- Yang et al., "SUPERB: Speech Processing Universal PERformance Benchmark," Interspeech 2021

> **Fri Sep 18 — Project proposal due (2 pages).**

08 · Tue Sep 22 — Prompting and parameter-efficient adaptation

- Lester, Al-Rfou & Constant, "The Power of Scale for Parameter-Efficient Prompt Tuning," EMNLP 2021
- Shi, Zhang & Gao, "Contrastive Learning for Speech Emotion Domain Generalization via Soft Prompt Tuning," Interspeech 2025

Unit III — Audio-language models and emotion

General hearing abilities bolted onto language models — and what it takes to make them hear affect.

09 · Thu Sep 24 — Audio-language models I

- Deshmukh, Elizalde, Singh & Wang, "Pengi: An Audio Language Model for Audio Tasks," NeurIPS 2023
- Tang et al., "SALMONN: Towards Generic Hearing Abilities for Large Language Models," ICLR 2024

10 · Tue Sep 29 — Audio-language models II

- Chu et al., "Qwen-Audio," arXiv:2311.07919, 2023
- Gong, Luo, Liu, Karlinsky & Glass, "Listen, Think, and Understand," ICLR 2024

11 · Thu Oct 1 — Making audio LLMs emotion-aware

- Du, Lu, Zhou & Gao, "Emotion-Aware Audio Large Language Models with Dual Cross-Attention and Context-Aware Instruction Tuning," Interspeech 2025

12 · Tue Oct 6 — Structured prompting and low-resource affect

- Shi, Du, Hong & Gao, "Plug-and-Play Emotion Graphs for Compositional Prompting in Zero-Shot Speech Emotion Recognition," IEEE ICASSP 2026
- Du, Shi, Myerston, Lu, Zhou & Gao, "Role-Guided Annotation and Prototype-Aligned Representation Learning for Historical Literature Sentiment Classification," Findings of EMNLP 2025

> **Thu Oct 8 — No class. Fall Break, October 8–11.**

13 · Tue Oct 13 — Adapting at test time

- Wang, Shelhamer, Liu, Olshausen & Darrell, "Tent: Fully Test-Time Adaptation by Entropy Minimization," ICLR 2021
- Shi, Du, Hong & Gao, "EMO-TTA: Improving Test-Time Adaptation of Audio-Language Models for Speech Emotion Recognition," IEEE ICASSP 2026

Unit IV — Generating emotional speech

Codecs, zero-shot synthesis, style control, and aligning generation with human preference.

14 · Thu Oct 15 — Neural audio codecs

- Zeghidour, Luebs, Omran, Skoglund & Tagliasacchi, "SoundStream," *IEEE/ACM TASLP* 2021
- Défossez, Copet, Synnaeve & Adi, "High Fidelity Neural Audio Compression (EnCodec)," arXiv:2210.13438, 2022

15 · Tue Oct 20 — Codecs that preserve affect

- Shi, Du, Song, Hong, Zhang & Gao, "AffectCodec: Emotion-Preserving Neural Speech Codec for Expressive Speech Modeling," Findings of ACL 2026
- Borsos et al., "AudioLM: A Language Modeling Approach to Audio Generation," *IEEE/ACM TASLP* 2023

16 · Thu Oct 22 — Zero-shot text-to-speech

- Wang et al., "Neural Codec Language Models Are Zero-Shot Text to Speech Synthesizers (VALL-E)," arXiv:2301.02111, 2023
- Shen et al., "NaturalSpeech 2," ICLR 2024

17 · Tue Oct 27 — Style and expressivity

- Li, Han, Raghavan, Mischler & Mesgarani, "StyleTTS 2," NeurIPS 2023

> **Mon Oct 26 — Last day to withdraw.**

18 · Thu Oct 29 — Aligning generative models with preferences

- Rafailov, Sharma, Mitchell, Manning, Ermon & Finn, "Direct Preference Optimization," NeurIPS 2023
- Wallace et al., "Diffusion Model Alignment Using Direct Preference Optimization," CVPR 2024

> **Fri Oct 30 — Midterm progress report due (4 pages).**
> **Tue Nov 3 — No class. Election Day.**

19 · Thu Nov 5 — Preference optimization for emotional TTS

- Shi, Du, Song, Hong, Zhang & Gao, "Emo-BPO: Emotion Bidirectional Preference Optimization for Diffusion-based Emotional TTS," Interspeech 2026
- Shi, Du, He, Hong & Gao, "Emotion-Aligned Generation in Diffusion Text-to-Speech Models via Preference-Guided Optimization," IEEE ICASSP 2026

Unit V — Interpretability and robustness

What the features mean, and what happens to them outside the training distribution.

20 · Tue Nov 10 — Opening the black box: interpretable emotion control

- Huben, Cunningham, Smith, Ewart & Sharkey, "Sparse Autoencoders Find Highly Interpretable Features in Language Models," ICLR 2024
- Du, Shi, Lu, Zhou & Gao, "Sparse Autoencoders for Interpretable Emotion Control in Text-to-Speech," ICML 2026

21 · Thu Nov 12 — Domain adaptation

- Ganin et al., "Domain-Adversarial Training of Neural Networks," *JMLR* 17(59):1–35, 2016
- Tzeng, Hoffman, Saenko & Darrell, "Adversarial Discriminative Domain Adaptation," CVPR 2017

22 · Tue Nov 17 — Distribution shift in deployed affect systems

- Lee, Lee, Lee & Shin, "A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks," NeurIPS 2018
- Gao, Baucom, Gordon, Rose, Wang & Stankovic, "E-ADDA: Unsupervised Adversarial Domain Adaptation Enhanced by a New Mahalanobis Distance Loss," IEEE SmartComp 2023

23 · Thu Nov 19 — Affect sensing in the wild: health deployments

- Gao, Salekin, Gordon, Rose, Wang & Stankovic, "Emotion Recognition Robust to Indoor Environmental Distortions Using Out-of-Distribution Detection," *ACM Transactions on Computing for Healthcare*, 2021
- Koltermann, Clapham, Blackwell, Jung, Burnet, Gao, Shao et al., "Gait-Guard: Turn-aware Freezing of Gait Detection for Non-intrusive Intervention Systems," IEEE/ACM CHASE 2024

24 · Tue Nov 24 — Beyond speech: multimodal and physiological affect *(online)*

November 23–24 are university-designated remote instruction days; this meeting is held online.

- Tsai, Bai, Liang, Kolter, Morency & Salakhutdinov, "Multimodal Transformer for Unaligned Multimodal Language Sequences," ACL 2019
- Mersa, Jiang, Gao, Li & Zhang, "Hierarchical Convolution Multibranch Transformer for EEG Signals," IEEE ICASSP 2026

> **Thu Nov 26 — No class. Thanksgiving Break, November 25–29.**

Unit VI — Project presentations

No assigned reading. The papers this week are yours.

25 · Tue Dec 1 — Final project presentations, session I

26 · Thu Dec 3 — Final project presentations, session II

> **Fri Dec 11 — Final report due (8 pages, conference format). No final exam.**

December 4 is the last day of classes. The university exam period runs December 7–11 and 14–15.

7. Supplementary reading

Optional. Useful for project work and for filling gaps in background.

- Zadeh, Chen, Poria, Cambria & Morency, "Tensor Fusion Network for Multimodal Sentiment Analysis," EMNLP 2017, pp. 1103–1114
- Bagher Zadeh, Liang, Poria, Cambria & Morency, "Multimodal Language Analysis in the Wild: CMU-MOSEI Dataset and Interpretable Dynamic Fusion Graph," ACL 2018, pp. 2236–2246
- Liang, Li & Srikant, "Enhancing the Reliability of Out-of-Distribution Image Detection in Neural Networks (ODIN)," ICLR 2018
- Kong, Goel, Badlani, Ping, Valle & Catanzaro, "Audio Flamingo," ICML 2024
- Macdonald, Chu, Stankovic, Zhou, Shao & Gao, "MiddleGAN: Generating Inter-domain Images with GANs," *TMLR* 2024
- Nair, Nafee, Jiang, Gao, Chen & Zhang, "Confidence-Aware Ranker Ensembles for Robust In-Context Knowledge Editing," Findings of ACL 2026
- Clapham, Zhou, MacDonald, Koltermann, Gao & Shao, "ElectroMeter: The Practical Electrolyte Measurement System," IEEE/ACM CHASE 2025
- Ko, Gao, Wang, Wijayasingha, Wright, Gordon, Wang, Stankovic & Rose, "Challenges to Recruiting Dementia Caregiving Dyads in Community-Based Settings," *JMIR Formative Research*, 2025

8. Policies

Short version: come to class, do your own thinking, tell me early when something goes wrong.

Attendance. This is a seminar; attendance *is* the course. More than two unexcused absences will materially affect your participation grade. If you must miss a meeting, notify the instructor in advance. Absences for illness, religious observance, or documented emergency are excused.

Late work. Reading responses are not accepted late — the point is that you have read the paper before we discuss it, and the two dropped scores exist to absorb ordinary life. Project deliverables may be submitted up to 48 hours late with a 10% penalty; beyond that, only by prior arrangement. *Prior arrangement means prior:* an extension asked for on Wednesday is a conversation; the same request on Saturday, after a Friday deadline, is a different one. Ask early.

Generative AI. You *may* use AI tools for coding assistance and debugging, and for language polishing on written work. You *may not* use them for generating your reading responses, your critique of a paper, or the substantive content of your project report — those are the thinking this course is designed to teach. Any substantial AI assistance must be disclosed in a brief note at the end of the submission; undisclosed use is an honor code matter. Note also that AI tools fabricate citations with some regularity — you are responsible for every reference you submit.

Honor Code. All work is governed by the William & Mary Honor Code. Collaboration on project work is encouraged within your declared team; representing others' work or ideas as your own is not permitted. When in doubt about whether a form of collaboration is acceptable, ask before rather than after.

Accessibility. William & Mary accommodates students with disabilities. Contact Student Accessibility Services to establish eligibility, then share the accommodation letter with the instructor as early in the semester as possible — arrangements need lead time to be useful.

Wellbeing. Graduate study is demanding, and a seminar that meets twice weekly makes it hard to disappear quietly during a difficult stretch. If something is going on, tell me early. Nearly every problem is easier to solve in September than in November. The W&M Counseling Center is available to students at no cost.

Research ethics. Several projects in this course will involve human affective data. Any project collecting new data from human participants requires **IRB review before collection begins** — plan for that timeline in your proposal. Projects using existing corpora must respect the license and use restrictions of those corpora; several of the datasets on the reading list restrict redistribution and commercial use.

9. Key dates at a glance

Date	Event
Thu, Aug 27	First meeting. Presentation sign-ups.
Fri, Sep 18	Project proposal due (2 pages)
Oct 8–11	Fall Break — no class October 8
Mon, Oct 26	Last day to withdraw
Fri, Oct 30	Midterm progress report due (4 pages)
Tue, Nov 3	Election Day — no class
Tue, Nov 24	Remote instruction day — meeting held online
Nov 25–29	Thanksgiving Break — no class November 26
Dec 1 & 3	Final project presentations
Fri, Dec 4	Last day of classes
Fri, Dec 11	Final report due (8 pages)

CSCI 680 · Affective Computing · William & Mary · Fall 2026. Syllabus subject to change. Changes are announced in class and posted to the course website.