Gesture–Discourse Alignment Project

Research Prospectus (Public Repository)

Author: Gefei Liu
Status: Independent Research / PhD Application Portfolio
Stack: Python (YOLOv8, MediaPipe) & R (reticulate, tidyverse)

Purpose

This prospectus documents the research motivation, conceptual framing, and analytic logic underlying a publicly available computational pipeline for studying gesture–discourse alignment in instructional settings.

1. Why This Project Is Public

This repository accompanies an independent research project developed in preparation for PhD applications.
Its purpose is **not** to present a finalized empirical paper but to transparently document the research question, methodological decisions, and analytic pipeline supporting an ongoing line of inquiry.

The repository is intended as a research artifact demonstrating how a theoretical question about gesture and discourse structure can be operationalized in a fully reproducible computational framework.

2. Research Motivation

In cognitively demanding instructional contexts, gesture may provide early temporal cues to the organization of discourse, particularly when linguistic processing is delayed. Prior work suggests that gestures can reflect aspects of instructional structure not fully explicit in speech, but existing approaches rely heavily on qualitative or categorical annotation, limiting scalability and temporal precision.

This project asks whether gesture can instead be modeled as a **continuous behavioral signal** whose intensity systematically aligns with **independently defined discourse-structural moments**. Rather than examining gesture meaning or type, the focus is on *when* gesture activity increases relative to instructional structure, enabling a quantitative, time-resolved analysis of gesture–discourse alignment in naturalistic classroom settings.

3. Conceptual Framework and Pipeline Overview

This project adopts a **temporal-structural** approach to gesture in instructional discourse and implements it through a fully reproducible, end-to-end analytic pipeline.

At the conceptual level, **gesture is treated as a continuous kinematic signal** rather than a symbolic or categorical unit. Hand movement is modeled as time-varying motion trajectories, from which salient gesture events are extracted as velocity-based peaks without categorization by type, form, or meaning. **Discourse structure is defined independently of gesture**, specified a priori based on instructional organization (e.g., topic shifts or phase boundaries), rather than inferred from gesture activity.

The primary analytic focus is **temporal alignment, not semantic correspondence**. Instead of asking what gestures represent, the framework asks *when* gesture intensity increases relative to independently defined structural moments. Structural relevance is therefore evaluated **statistically rather than categorically** by comparing gesture event rates inside versus outside structural time windows using rate-normalized metrics.

These conceptual commitments are realized through the following pipeline stages:

1. **Video preprocessing and localization**
   Classroom videos are processed to identify teacher presence and extract wrist trajectories using computer vision tools.
2. **Kinematic signal construction**
   Frame-level wrist positions are transformed into continuous velocity signals representing hand movement over time.
3. **Gesture event detection**
   Salient gesture events are identified via high-quantile thresholding and temporal clustering of velocity peaks.
4. **Discourse–gesture alignment**
   Detected gesture events are aligned to pre-annotated discourse-structural points using symmetric temporal windows.
5. **Statistical evaluation**
   Alignment is quantified using rate-normalized enrichment statistics and evaluated against permutation-based null models, including both global temporal nulls and teacher-restricted nulls to control for activity density and visibility.

Annotation confidence is treated as a **robustness dimension rather than a defining criterion**: confidence ratings on structural points are used solely for sensitivity analyses assessing result stability under stricter annotation assumptions. The framework remains **intent-agnostic and outcome-neutral**, making no claims about gesture function, communicative intent, or learning outcomes.

Intermediate outputs (CSV files and Excel alignment workbooks) are preserved throughout the pipeline to support transparency, auditability, and reproducibility at scale.

4. Current Status of the Project

The code in this repository reflects an active research implementation that has been piloted on multiple classroom lecture segments. Across the currently analyzed materials, gesture events show clear temporal enrichment around discourse-structural points, with observed effects consistently exceeding permutation-based baseline expectations.

Across all tested lecture segments, enrichment statistics are statistically significant under both global and teacher-restricted null models ($p < 0.05$) and remain stable across reasonable variations in detection thresholds and alignment windows.

These findings are presented as **proof-of-concept evidence** demonstrating that the proposed pipeline can reliably detect structure-linked temporal patterning in gesture data, rather than as final empirical claims about gesture function or pedagogy.

5. Scope and Limitations

This repository is intended to demonstrate:

- a methodological framework
- a reproducible analytic pipeline
- a theoretically motivated operationalization of gesture–structure alignment

It is **not** intended to:

- serve as a finalized empirical paper
- claim generalizable effect sizes
- replace detailed qualitative analysis of gesture meaning

The project is designed to be extended, refined, and potentially integrated with complementary analytic approaches.

6. Intended Audience

This prospectus and repository are intended for:

- Faculty reviewing PhD applications
- Potential advisors evaluating methodological fit

- Researchers interested in quantitative approaches to gesture and instruction

Readers are encouraged to treat this repository as a **research artifact**, not a polished product.

7. How to Cite or Reference

If you wish to reference this work, please cite the GitHub repository and note that the project represents **ongoing, unpublished research**.