Final Project Proposal

Do Men and Women Speak Differently About Monetary Policy?

A Text-as-Data Analysis of Hawkish and Dovish Language in Central Bank Speeches

1. Motivation and Research Questions

Central bank communication plays a key role in shaping market expectations and public trust. Yet, most research focuses on *what* central bankers say, not *how* their communication may differ across gender lines. This project investigates whether male and female central bankers systematically differ in their **tone, framing, and semantic use** when discussing monetary policy.

Specifically, the project asks:

1. Do male and female central bankers use systematically different language when discussing monetary policy decisions?
2. Are there gender differences in how *hawkish* or *dovish* concepts are framed and contextualized?
3. What *semantic neighborhoods* (meanings, connotations) surround key words like *inflation*, *growth*, or *stability* for male vs. female speakers?

This study contributes to debates on both **gender representation in economic policymaking** and **linguistic framing in monetary communication**, with implications for understanding diversity, credibility, and tone in central banking institutions.

2. Data

The project uses the **Central Bank Speeches (CBS)** dataset, which compiles public speeches from over 100 central banks worldwide, including detailed metadata.

- **Corpus:** ~14,000 speeches (in English or translated English versions) from 1996–2024.
- **Metadata variables:** speaker name, gender (derived from manual or automated tagging), institution, date, country, title, and text body.
- The sample will be limited to speeches in English, focusing on institutions with sufficient female speaker representation
- If time permits, additional contextual variables (e.g., inflation rate, unemployment rate) will be matched by country-year to control for macroeconomic conditions at the time of speech.

3. Methodology

Step 1. Preprocessing

- Tokenize, lowercase, and lemmatize texts.Remove stopwords, punctuation, and non-economic noise terms. Retain key economic vocabulary (e.g., “inflation”, “employment”, “policy”, “rates”).

Step 2. Hawkish–Dovish Dictionary Scoring

- Apply an existing **hawk–dove dictionary** (Apel & Blix Grimaldi 2012; Lucca & Trebbi 2009) that captures pro-tightening vs. pro-loosening language.
- Calculate a *hawkishness score* for each speech:
- Compare mean tone scores across **gender**, **institution**, and **year**.

Step 3. Word Embeddings and Semantic Context

To move beyond dictionary-based counts, I will train **Word embeddings** separately for male and female subsets of the corpus. This allows mapping the **semantic neighborhoods** of key monetary terms (e.g., *inflation*, *growth*, *inequality*).

Example:

- In male speeches: *inflation* → {“pressure”, “tightening”, “expectations”, “stability”}
- In female speeches: *inflation* → {“household”, “wages”, “inequality”, “consumers”}

To statistically validate these semantic differences, I will implement **embedding regression models**. This method quantifies how word meanings shift with covariates (e.g., gender, time, institution) rather than relying on subjective interpretation.

4. Expected Findings

1. Male speakers may emphasize **inflation control and monetary tightening** (hawkish).
2. Female speakers may emphasize **employment, welfare, and inclusive growth** (dovish, human-centered).
3. Word embeddings could show **semantic divergence** — e.g., “growth” associated with *output* vs. *equity* depending on gender.

5. Potential Limitations

- Fewer female central bankers → unbalanced corpus.
- Translation inconsistencies in multilingual sources.
- Word embedding comparisons sensitive to corpus size; mitigation via subsampling and validation checks.

6. References (Selected)

- *Lucca, D. O., & Trebbi, F. (2009). Measuring central bank communication: An automated approach. Journal of Monetary Economics, 56(2), 254–269. https://doi.org/10.1016/j.jmoneco.2008.12.007*
- *Masciandaro, D., Profeta, P., & Romelli, D. (2020). Yellen at Yellen: Gender and communication in central banking. Journal of Monetary Economics, 115, 178–196. https://doi.org/10.1016/j.jmoneco.2020.05.006*
- *Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2023). Embedding regression models for context-specific description and inference. American Political Science Review, 117(4), 1251–1270. https://doi.org/10.1017/S0003055423000304*
- *Rutkowska, A., & Szyszko, M. (2024). Dictionary-based sentiment analysis of monetary policy communication: On the applicability of lexicons. Quality & Quantity, 58(6), 5421–5444. https://doi.org/10.1007/s11135-024-01896-9*
- *Baranowski, P., Bennani, H., & Doryń, W. (2021). Do the ECB's introductory statements help predict monetary policy? Evidence from a tone analysis. European Journal of Political Economy, 66, Article 101967. https://doi.org/10.1016/j.ejpoleco.2020.101967*
- *Apel, M., & Grimaldi, M. B. (2012). The information content of central bank minutes (Sveriges Riksbank Working Paper Series No. 261; Riksbank Research Paper Series No. 92). Sveriges Riksbank. https://doi.org/10.2139/ssrn.2092575*
- *Hansen, S., & McMahon, M. (2016). Words versus deeds: How central banks manage expectations. VoxEU.org. https://cepr.org/voxeu/columns/words-versus-deeds-how-central-banks-manage-expectations*