An Auditable LLM Pipeline for Text Measurement: Conceptual Frame Profiling

Orsolya Putz · BME TMIT · AIME26, 12 Nov 2026

About me

dr. Putz, Orsolya

Cognitive Linguist

Cognitive Psychologist

NLP specialist

BME TMiT

Conceptual frames are the systems of world knowledge through which policymakers understand economic events, actions, and phenomena. Conceptual frame profiling (CFP) operationalizes this relation.

“The labor market is overheated.”

Sentiment tools: neutral (polarity 0.00)

CFP: the labor market is a malfunctioning machine



*Metaphors we live by** – even in the Federal Open Market Committee's (FOMC) minutes

Metaphorical expression

Inflation is climbing

Target domain

The abstract concept being described, e.g. inflation.

Source domain

The concrete domain through which a target is understood,e.g. vertical movement.

Conceptual metaphor

TARGET IS SOURCE – INFLATION IS VERTICAL MOVEMENT

1 in ~55 words

in the FOMC minutes is a metaphorical expression

32,154 expressions · 257 minutes · 1.76M words · 1992 – 2006

*Lakoff & Johnson (1980): metaphor is a mechanism of human thought that structures how we perceive the world and act within it

BME TMiT

Metaphors are glasses: they let us see the same concept from a different perspective

Target: INFLATION

WAR

“Inflation is an enemy.”

Implication

Enemies, damage, who is to blame

How to act?

fight against it, defeat it

ILLNESS

“Inflation is a disease.”

Implication

symptoms, spreading, relapse

How to act?

diagnose, treat, monitor

VERTICAL MOVEMENT

“Inflation remained elevated”

Implication

physical object moving up or down

How to act?

bring it down

FLOOD

“A wave of inflation.”

Implication

natural force beyond our control

How to act?

build dams, wait it out

The codebook: 8 source domains × 3 stain levels x 5 target domains

Frame	Strain level 1: control	Strain level 2: tension	Strain level 3: loss of control
Horizontal movement	move forward	roadblock	dead end
Vertical movement	fall, increase	skyrocket	plummet
Physical entity	solid	pressure	collapse
Liquid	flow	freeze	flood
Machine	function	overheat	breakdown
Living organism	growth	sick	dead
War / Game / Sport	strategy	fight	game over
Building	foundation	unstable	ruin

Target domains: Economy, Inflation, Financial, Market, Monetary Policy, and Labor Market

Metaphor identification methods

Human annotation

(MIPVU)

- Replicable, but labour-intensive
- Throughput: dozens of texts

Deep-learning detectors

(DeepMet, MelBERT)

- Binary label per token
- No source, target or state
- Not built to aggregate

Conceptual Frame Profiling

(this work)

- an LLM identifies the metaphorical expression + source + target + strain level + generates explanation
- Aggregates into document-level profiles

Workflow of Conceptual Frame Profiling (CFP)

Codebook

8 source domains

5 target domains

3 strain levels

›

LLM

Gemini 2.5 Pro reads the whole document, returns schema-bound JSON

›

3 runs

independent, temperature 1.0

›

Filter

keep expressions confirmed in ≥ 2 runs, fuzzy match

›

Profile

per document: shares of source and target domains; mean strain index = mean strain level per document

+ Validation: human annotation, another LLM

BME TMiT

Results

76.3%

expressions confirmed by an independent run

precision

84.5–92.8%

judged correct by 3 annotators

accuracy, 50 passages

κ 0.82–0.91

source-domain agreement between annotators

human reference

r = 0.89

open-weights Gemma 4 reproduces the profiles

model independence

Humans reviewed 15,000 words = 0.85% of the corpus, they were not blind, so recall is an upper estimate.

BME TMiT

Sentiment vs CFP : overlap in intensity, independence in structure

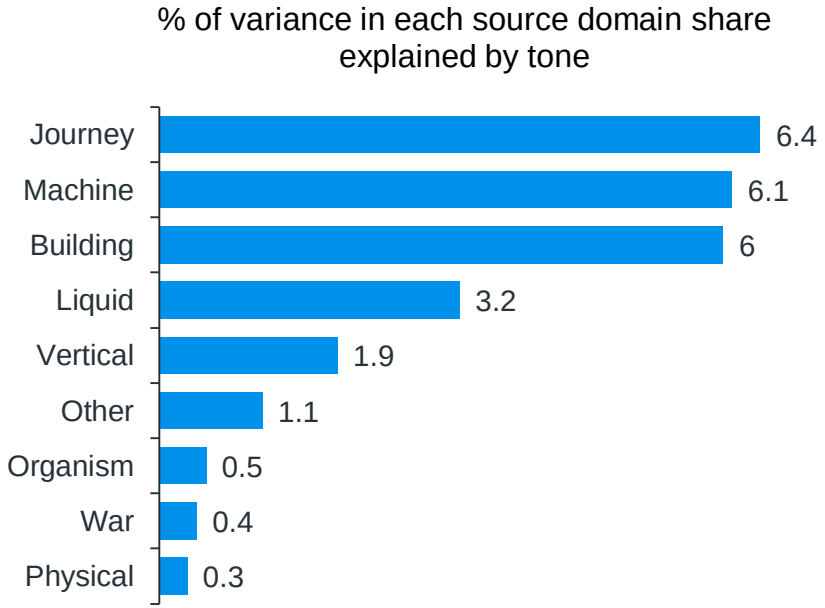
r = −0.66

strain index vs Loughran–McDonald; −0.52 vs FinBERT

Tone explains 44% of the strain index

but ≤ 6.4% of any source domain share

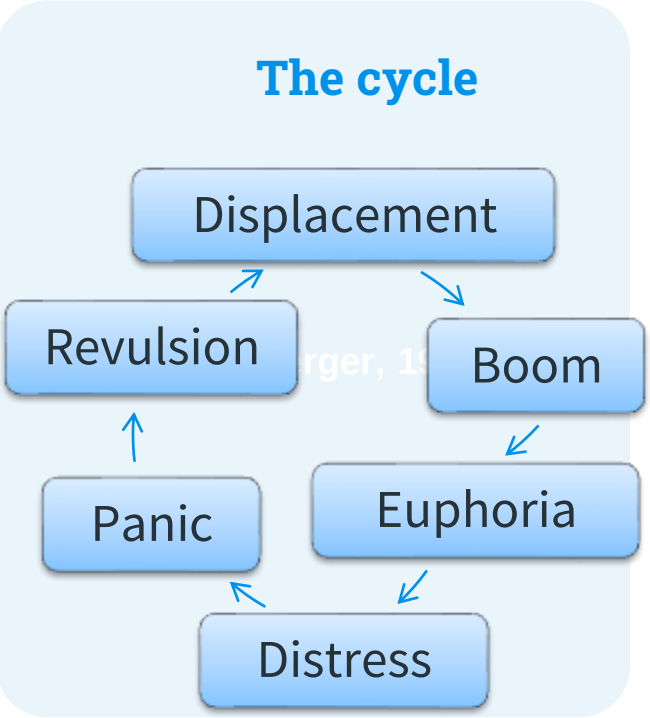
n = 257 minutes



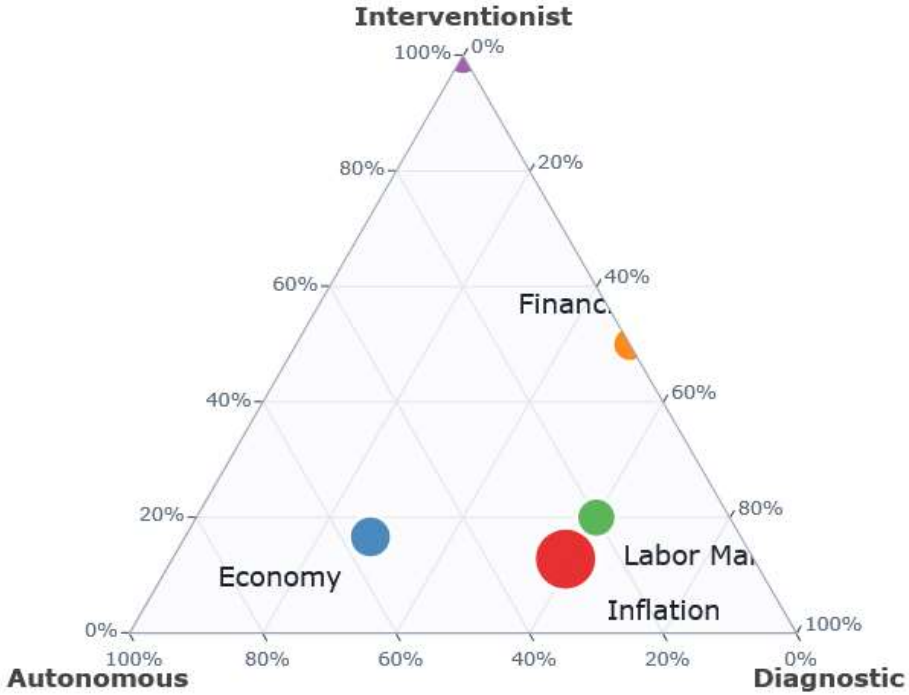
Sentiment vs CFP: the measures see different phases of a financial crisis (Kindleberger, 1978)

Phase vs Normal	Strain index	L–M polarity	FinBERT
Displacement	lower**	higher**	higher*
Boom	n.s.	n.s.	n.s.
Euphoria	lower**	n.s.	n.s.
Distress	higher*	lower**	lower**
Panic	higher*	lower**	lower**
Revulsion	higher*	lower**	lower**
Post-2022	higher*	lower**	lower**

Phases hand-coded (Kindleberger); * pFDR<0.05, ** pFDR<0.01; vs Normal, n = 99



CFP extended: new features – narrative stance

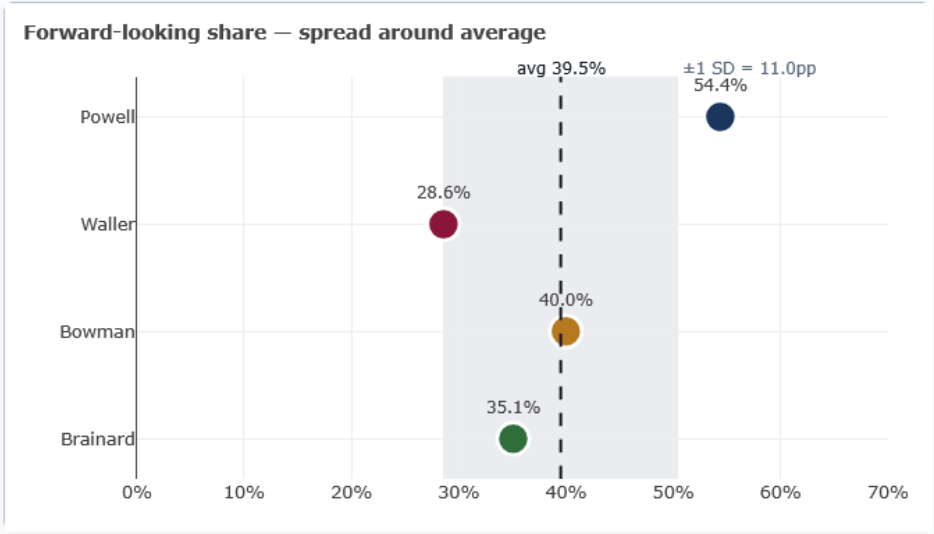


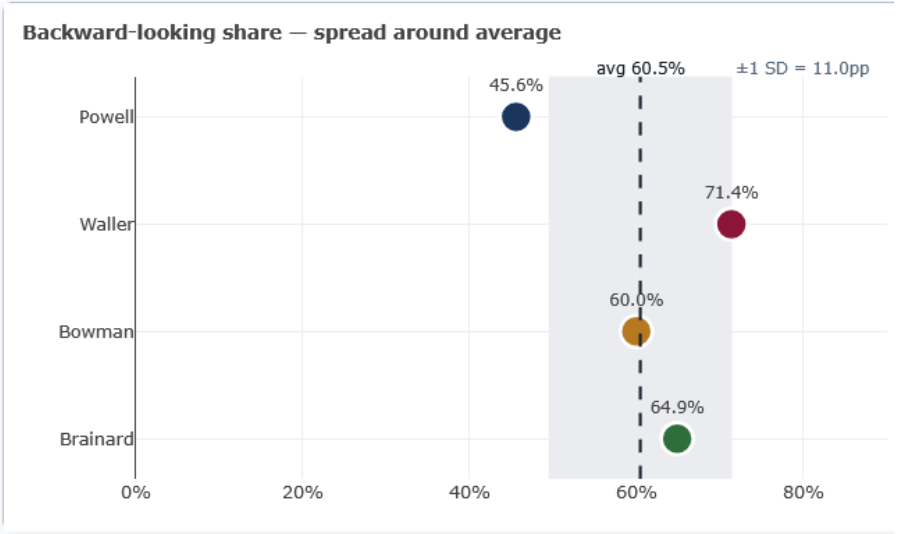
STANCE

The primary actor driving economic conditions: *Interventionist* (the Fed acts), *Autonomous* (the market moves on its own), or *Diagnostic* (static observation).

PILOT · 4 speeches (Powell, Waller, Bowman, Brainard) analyzed from 2022 ·

CFP extended: new features – time orientation





Time orientation: *Forward-looking* (intent, projection, and guidance) versus *Backward-looking* (track record and past data).

Future work

1 · Other jargons

Test CFP on other text types (e.g. political speeches, social media posts).

2 · The dot plot

Do the dimensions of a FOMC member's speech predict where the speaker sits relative to the committee?

3 · A fine-tuned LLM

Release an open model trained for CFP.

CFP core features (source · target · strain level) + features harmonized with the jargon

e.g. narrative stance and time orientation in FOMC members' speeches

Summary

1

"Metaphors we live by" – even in the driest texts.

CFP makes the conceptual frame of a document measurable.

2

CFP is feasible at corpus scale, and auditable.

Three runs, human review, a second LLM.

3

CFP measures what tone cannot see.

Frame composition is nearly independent of sentiment, and the profile can grow new dimensions.

Thank you for your attention.

putz.orsolya@tmit.bme.hu