

# Measuring Semantic Flow Without Direction: A Rhizomatic Protocol for Stereotype Translation in Cross-Cultural Language Technology

Gustavo Aviña Cerecer · Universidad Autónoma de San Luis Potosí (UASLP) · ✉ [gac@uaslp.mx](mailto:gac@uaslp.mx)  
 Gradients of Alterity (GA) · Reaffirm / De-signify / Escape (RDE) protocol



## WHAT PROBLEM DOES SEMANTIC FLOW WITHOUT DIRECTION SOLVE? MONOTRANSLATIONISM — all monotranslation is an epistemicide of human multiplicity

Language technology today translates the world from a single point: **English** — one grammar, one syntax, one alphabet, one ontology. Stereotype-correction benchmarks impose Western North American culture as the universal pole — in its libertarian ethics or in its woke ethics, it is still **only North American**. Every monotranslation flattens lexical-semantic diversity and produces **epistemicides**. The answer is not to reduce difference to a single corrective direction, but to **measure it, know it, and preserve it**. That is what this protocol does.

Every stereotype is, at bottom, one of two movements: a **culturocentric reaffirmation** — of one's own territory or of a foreign one — or an attempt to **escape culture itself**. That is why stereotype translation must be measured as semantic movement (Reaffirm / De-signify / Escape), never corrected toward a single cultural pole.

### 1 WHY DIRECTION-AGNOSTIC?

LLMs are deployed across linguistic and cultural contexts never represented in their training distributions, yet stereotype evaluation remains mostly **English-centric** and assumes a **single normative axis of correction**.

- It cannot distinguish a user who reproduces *their own* territorial codes (a Russian conservative endorsing Russian discourse) from one who reproduces *foreign* codes against their own (the Mexican *malinchista*, the diasporic subject).
- Collapsing bias into one axis with a "correct" pole confuses **measurement with prescription**.

C1 Direction-agnostic RDE protocol

C2 Abstract-machine detector: 526 patterns · 8 languages

C3 Refined velocity for short & long texts

C4 Deployed open-source web prototype + corporate calibration

**Our question:** across the same set of semantic operators, how does the user's translation relate to the provocateur discourse and to their own calibrated profile? The system measures *movement*; the user — or the deploying organization — interprets *meaning*.

### 2 RHIZOME AGAINST THE TREE

Chomsky's arborescent model derives meaning hierarchically from a root S by dichotomy. Deleuze & Guattari (1987) refuse this: any point of a rhizome can — and must — connect to anything other. Three implications: semantic territories are **heterogeneous**; identity and alterity are **effects of connection**; measurement must be **horizontal** (trace connections, not depth).

#### Molar lines → R

Well-defined segments coded by dispositives of power (family, school, state, army). Totalizers, naturalizers, binary exclusions, we/them splits. High transversal reach = high molar weight.

#### Molecular lines → D

Flows by thresholds, becoming; relations of speed and slowness. Modulators (*sometimes, depending on*), partial markers, gradulators. They soften the territory without leaving it.

#### Lines of flight → E

Not segmentary but abstract: signifiers escape the entire molar-molecular chain (a trans subject displacing binary discourse onto non-binary semantics; *ñuke mapu* against elite discourse).

### 3 THE RDE SCORES & GRADIENT

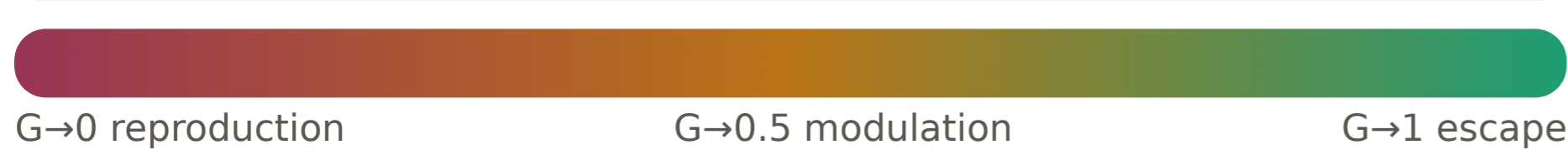
Given operator intensities  $I_M, I_m, I_F$  (molar, molecular, flight) in the interpretation and token overlap  $\omega$  with the provocateur:

$$R = \min\{1, 0.7 \cdot I_M / (I_M + I_m + I_F) + 0.3 \cdot \omega\}$$

$$D = I_m / (I_M + I_m + I_F)$$

$$E = \min\{1, 0.7 \cdot I_F / \Sigma + 0.3 \cdot (1 - \omega) \cdot I_F / \Sigma\}$$

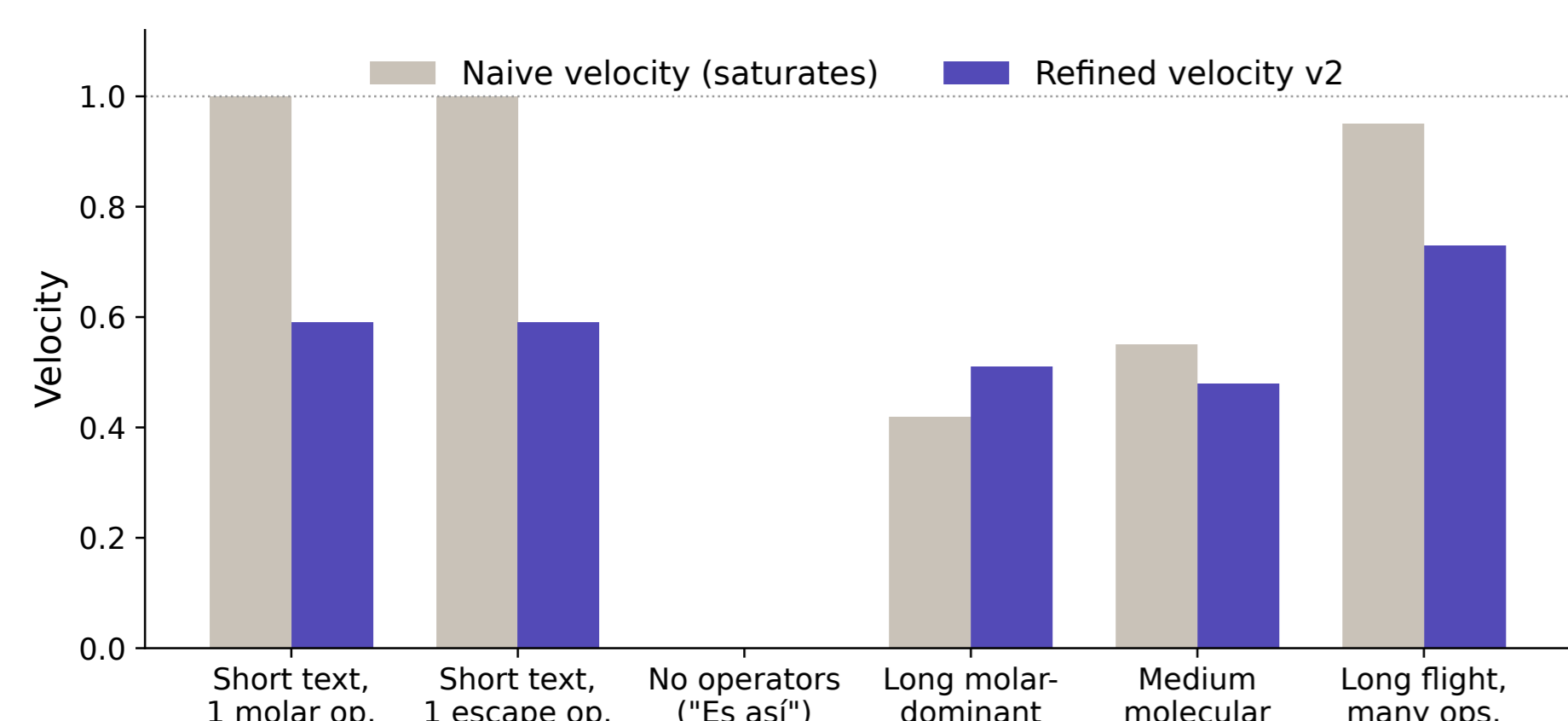
$$G = (R \cdot 0 + D \cdot 0.5 + E \cdot 1.0) / (R + D + E) \in [0, 1]$$



**No score is a priori better.** A Russian conservative reaffirming Russian discourse and a *malinchista* reaffirming Anglo upper-class discourse both score high R — the protocol reports *that* they reaffirm, not *which territory* they reaffirm toward. Chain provenance disambiguates assimilationist escape from cosmological flight.

### 4 REFINED VELOCITY FOR SHORT TEXTS

Naive velocity (edges per length) saturates on short interpretations. The refined metric combines **operator-subtype turnover**, **signifier diversity**, **machine alternation** across linear order, and **logarithmic density**.



Verified behavior of velocity v2: one-clause texts yield discriminative 0.4–0.7 values instead of saturating at 1.0; texts without operators correctly score 0.

### 5 SYSTEM ARCHITECTURE

#### Calibration

6-question survey: country of origin, migratory trajectory, social-epistemic position, languages of thought, position in power systems, theme(s) of alterity — relative positions, never identity categories.

#### Provocateur retrieval

Geopolitically meaningful matching: a Ukrainian user selecting sexism receives a Russian traditional-family discourse, not a generic Anglo one; a Mapuche user receives Chilean elite discourse.

#### Operator detection — Layer A + Layer B

A: 526 transparent, auditable linguistic patterns across 8 languages, grouped by abstract machine. B: multilingual embeddings (LaBSE, XLM-R) fill gaps by similarity to seed-corpus centroids.

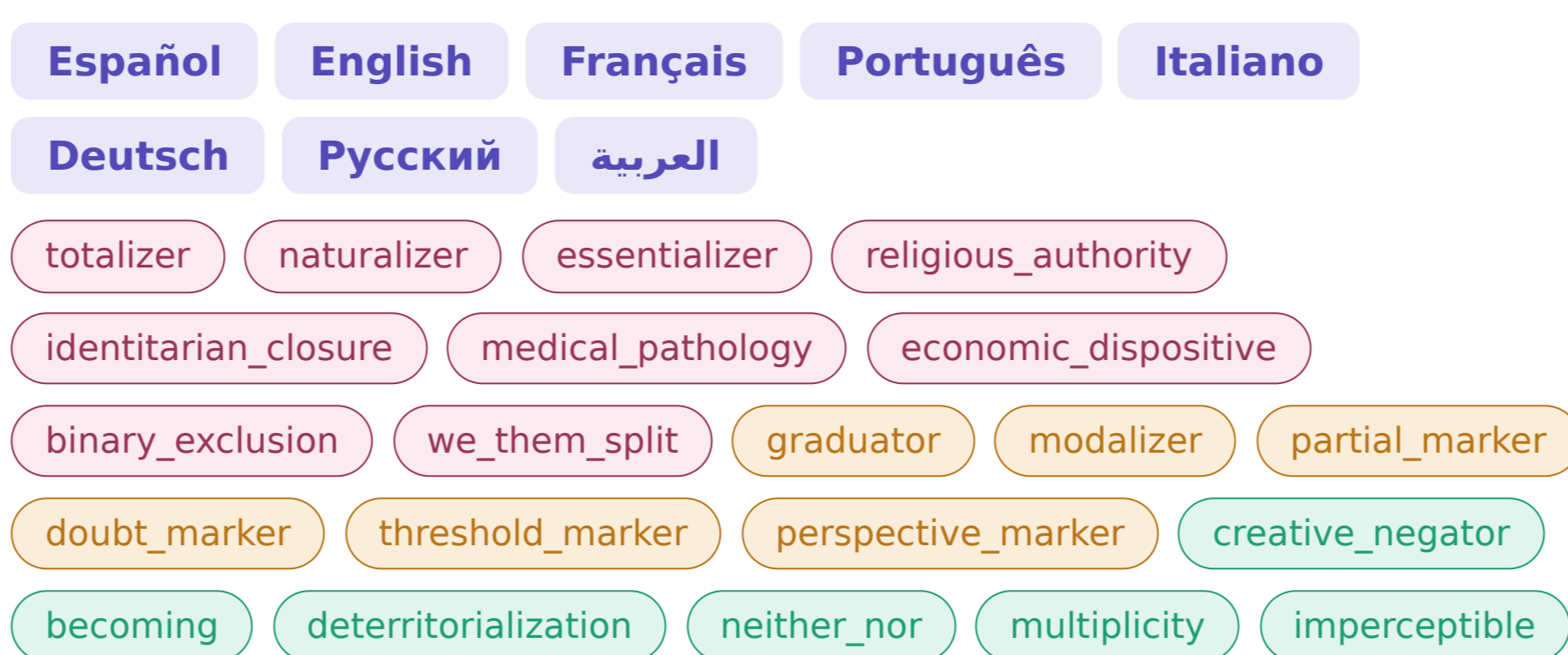
#### R / D / E + composite gradient

Three independent scores plus G; interpretation rendered with operators highlighted in machine-specific colors.

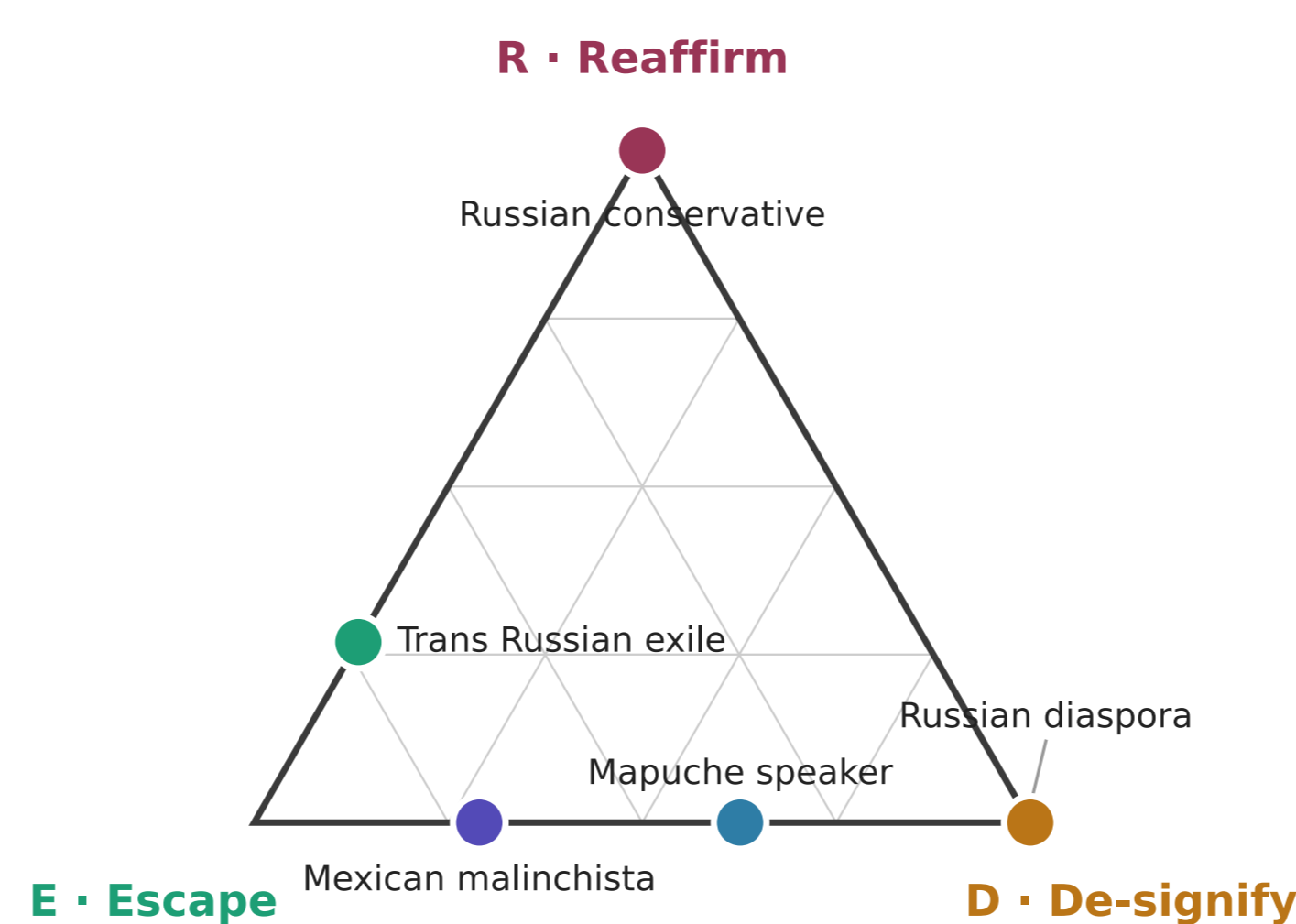
#### Tier-gated dashboard

Free → Pro → Admin/Corporate.

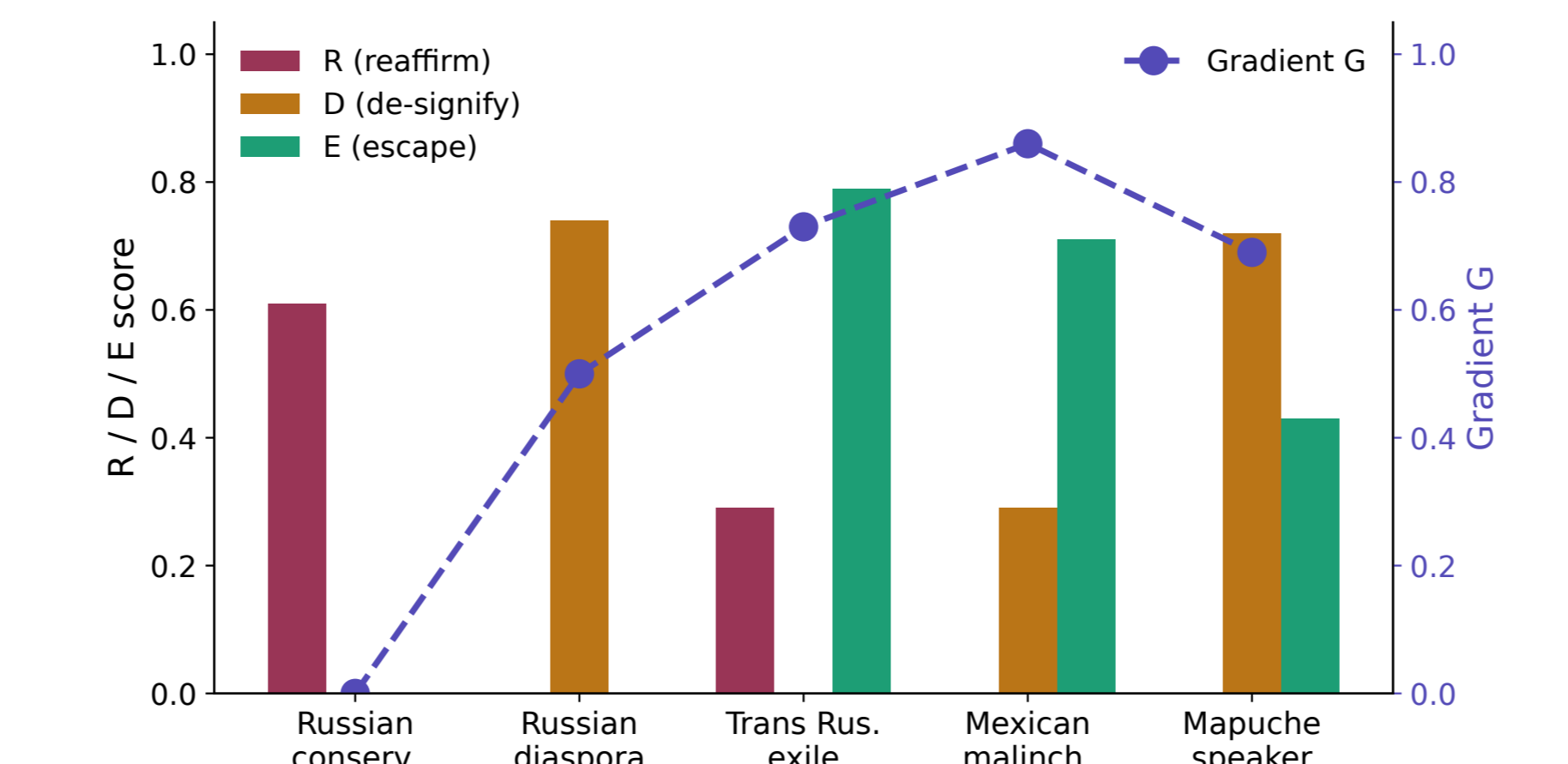
### 6 526 PATTERNS · 8 LANGUAGES



### 7 FIVE EXTREME PROFILES



Profile	R	D	E	G
Russian conservative	0.61	0.00	0.00	0.00
Russian diaspora	0.00	0.74	0.00	0.50
Trans Russian exile	0.29	0.00	0.79	0.73
Mexican <i>malinchista</i>	0.00	0.29	0.71	0.86
Mapuche speaker	0.00	0.72	0.43	0.69



Coherent and distinct RDE signatures. Chain provenance — not the scalar gradient alone — separates assimilationist E (*malinchista*: escape lands in the provocateur's adjacent territory) from cosmological-flight E (Mapuche: *ñuke mapu* originates in a distinct semiotic regime). Same gradient class, different navigational meaning.

### 8 DEPLOYMENT TIERS

#### Free — individual reflective use

Composite gradient, R/D/E breakdown, interpretation highlighted in machine colors, plain-language summary, recalibration and comparison.

#### Pro — researchers

Inspect the provocateur's own RDE decomposition, manipulate gradient sliders, retrieve real-world discourses matching the manipulated profile.

#### Admin / Corporate — tolerance bands

Organizations whose personnel translate sensitive cross-cultural material (diplomatic services, NGOs, content moderation, journalism) define **tolerance bands per axis** — e.g. an aid organization in a conflict zone: low R, high D, moderate E. Descriptive of how the team's semantic flow distributes — never prescriptive of "correct" translations. A single global threshold for acceptable stereotype neutralization is itself a colonizing move.

### 9 ADVANCES TO DATE

POST-SUBMISSION · JUNE 2026

- Serverless production deployment.** The prototype now runs as a static single-file frontend plus an edge worker proxying the analysis API with protected credentials and CORS — zero-cost free tier (~100k requests/day), no cold starts.
- Corpus-grounded signifier weights.** A new `/api/ngrams` endpoint weights detected signifiers with real Google Books Ngram frequencies (Spanish corpus 2000–2019), anchoring molar weight in measured transversal presence — e.g. *familia* ≈ 3.7 × 10<sup>-7</sup> vs *sexualidad* ≈ 1.4 × 10<sup>-8</sup>.
- Dynamic provocateur generation.** Beyond the curated corpus, provocateurs are now generated per calibrated profile by an LLM under the geopolitical-relevance constraint, with content warnings preserved.
- Extended polarity gradient.** A signed variant GA = -1.0·R + 0.5·D + 1.0·E ∈ [-2, +2] complements G, making territorial anchoring vs. active flight directly readable on one axis.
- Detector v2 verified across all 8 languages**, including Arabic syntax (مقدسة) and German compounds (*natürliche Ordnung*); connection graph filtered to the top-30 edges (intensity ≥ 0.5) for legible visualization.
- UX for the translation industry.** Calibration questions rewritten with domain feedback (epistemology vs. episteme distinction corrected); every number on the dashboard now ships with a plain-language explanation; accessible earth-tone palette.
- Computing Multiplicity platform.** The instrument — now the *Semantic Symbiont (Gradients of Alterity)* — is being integrated with a live *Observatory of Monolinguisism* tracking language extinction and epistemic loss, targeting the international translation and interpretation industry.

### 10 EVALUATION

- 5 user profiles + battery of **16 short/long texts × 8 languages**, each constructed as predominantly molar, molecular, flight, or mixed — classified consistently with expectation, no tuning between runs.
- Ñuke mapu* registers as flight by **semiotic-regime provenance** (indigenous cosmology vs. Chilean elite molar foundation) — not by any vocabulary list.
- Complementary to model-output benchmarks: it captures **what users do** with stereotype material.

### 11 LIMITATIONS & ETHICS

- Hand-curated patterns reflect the authors' linguistic competence; coverage is partial. Biased embedding seeds → biased machine attribution.
- Irony/quotation of molar operators is not distinguished from sincere use; movement ≠ understanding.
- Tolerance bands must **never** become individual performance metrics — that would reproduce the segmental, normalizing control D&G critique. No retention without consent; no identity-based demographic questions; content warnings precede every provocateur.
- Large-scale validation with diverse human users is necessary future work.

### 12 KEY REFERENCES

- Deleuze, G. & Guattari, F. (1987). *A Thousand Plateaus*. Univ. of Minnesota Press.  
 Chomsky, N. (1957). *Syntactic Structures*; (1965). *Aspects of the Theory of Syntax*.  
 Nadeem, M. et al. (2021). StereoSet. *ACL-IJCNLP*. · Nangia, N. et al. (2020). CrowS-Pairs. *EMNLP*.  
 Parrish, A. et al. (2022). BBQ. *Findings of ACL*. · Blodgett, S.L. et al. (2021). Stereotyping Norwegian salmon. *ACL*.  
 Durmus, E. et al. (2023). Global opinions in LMs. *arXiv:2306.16388*. · Chiu, Y.Y. et al. (2024). CulturalBench. · Li, C. et al. (2024). CulturePark. *NeurIPS*.  
 Amoreo, L. (2020). *Cloud Ethics*. Duke UP. · Parisi, L. (2019). The alien subject of AI. *Subjectivity* 12(1).