

# DiaLLM: An Investigation into the Robustness–Generation Gap in English Dialect Adaptation

Jordan Painter<sup>1</sup>, Dipankar Srirag<sup>3</sup>, Adarsh Kappiyath<sup>1</sup>, Diptesh Kanojia<sup>1,2</sup>, Aditya Joshi<sup>3</sup>, Lu Yin<sup>1</sup>

<sup>1</sup>NICE Research Group and <sup>2</sup>People-Centred AI, University of Surrey, UK <sup>3</sup>University of New South Wales, Sydney, Australia



## Motivation

LLMs increasingly understand dialectal English, yet still produce standard, US-leaning English.

### Three users chasing a late parcel:

*Indian English:* My parcel is not yet coming; kindly **do the needful** and **prepone** the delivery.

*Northern British English:* **Me** parcel's not arrived; **summat** were up **wi'** the courier.

*Australian English:* Me parcel still hasn't turned up, **reckon** it's good this **arvo**, eh?

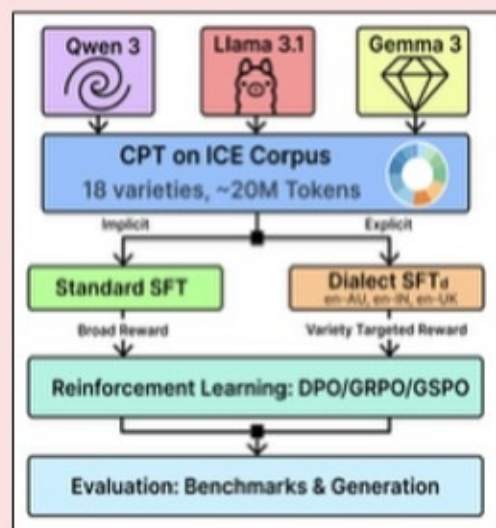
A modern LLM *understands* all three, yet answers each in flattened Standard American English.

A model that understands every variety but speaks only one has solved half the problem.

Prior dialect methods (TADA, HyperLoRA, LoRDD) target robustness to dialectal *input*.

DiaLLM asks whether a model can be adapted to *produce* a variety, and which pipeline stage makes that happen.

## The DiaLLM Pipeline



We continually pretrain Llama-3.1-8B, Qwen3-8B and Gemma-3-4B-IT on the International Corpus of English (18 varieties, ~20M tokens). Two threads follow from each checkpoint. The **implicit** thread fine-tunes on standard preference data and aligns broadly. The **explicit** thread fine-tunes on data converted into one variety with Multi-VALUE (en-AU, en-IN or en-UK) and aligns towards that variety.

Both threads are aligned with DPO, GRPO and GSPO. GRPO and GSPO optimise a reward that weights eWAVE feature density against meaning preservation:

$$R = \lambda \cdot \phi_{\text{dial}} + \frac{1}{2}(1-\lambda) \cdot (\phi_{\text{comet}} + \phi_{\text{cos}}), \quad \lambda = 0.8$$

$\phi_{\text{dial}}$ : classifier over 135 eWAVE features.  $\phi_{\text{comet}}$ ,  $\phi_{\text{cos}}$ : meaning preservation against the SFT answer.

We evaluate on six benchmarks (BBH, GPQA, GLUE, VALUE, BESSTIE, DialectBench) and on 25 open prompts judged by six native or near-native annotators and by Phi-4.

## ① Can a Model Speak a Variety?

Yes. Fine-tuning on dialect-specific data produces output that speakers recognise as dialectal.

Annotators picked the dialect-specific model over the instruct model (Llama-3.1-8B)

84%

en-IN

87%

en-UK

48%

en-AU

Australian English is the exception: annotators show no clear preference, consistent with its subtler features. (Table 4, T1; ties excluded)

Same prompt, broad vs. variety-targeted model (selected examples, Table 3)

en-UK · Qwen3-8B · GRPO

Prompt: What's the best piece of advice someone's ever given you?

Broad model

"Be kind to others, and treat them the way you would like to be treated." ... a fundamental principle of many cultures.

Features: possessive me; first-person singular us

Variety-targeted model

This advice has helped us navigate through difficult situations and make decisions that align with **me** core values.

en-AU · Llama-3.1-8B · DPO

Prompt: What are some small things that can brighten up a rough day?

Broad model

... small things, depending on what makes you feel happy and energized... Treat yourself to a small indulgence...

Features: like as a discourse particle; yourself

Variety-targeted model

... what makes you feel like happy... Kindness to **yourself**: Treat **yourself** to a small indulgence...

en-IN · Gemma-3-4B-IT · GSPO

Prompt: What's your take on people who are always late?

Broad model

... I understand that being late can be a frustrating experience for others... it's not always appropriate to judge someone solely based on their punctuality...

Features: stative progressive; article and progressive overextension

Variety-targeted model

... I **am understanding** that being late can be frustrating, especially where the punctuality is highly valued... it's not appropriate to judge or **are criticizing** them without knowing the full contexts...

## ③ More Reward ≠ More Dialect

The method the eWAVE reward likes most, GRPO, is never the one human or LLM judges pick as most dialectal.

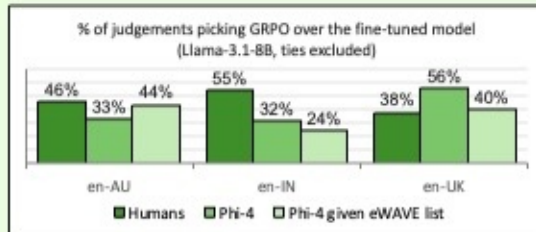
By the eWAVE reward

GRPO ▲ most dialect features, in all three families  
▼ fewer dialect features than before alignment

Judged most dialectal

DPO en-IN and en-UK (humans)  
GSPO en-AU  
GRPO never the most-picked

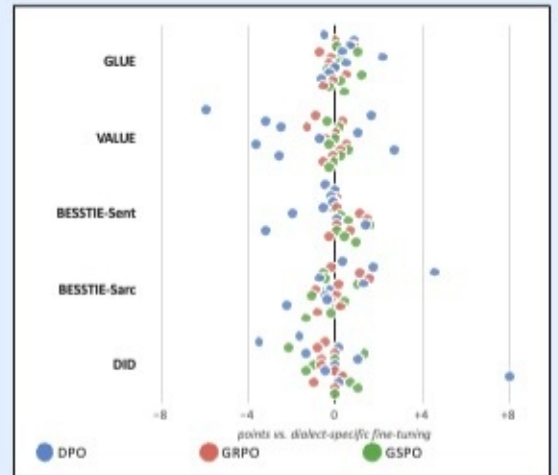
Against the fine-tuned model it builds on, GRPO shows no consistent advantage with any judge



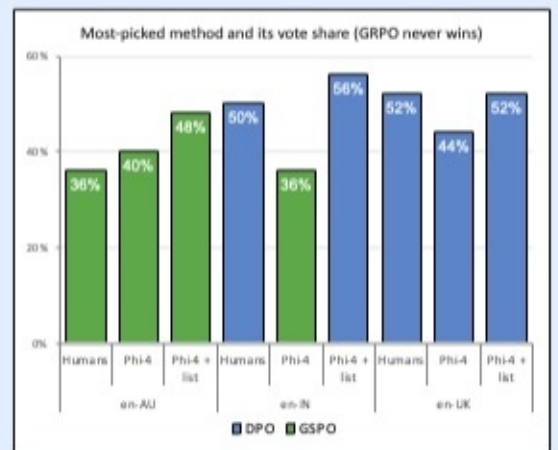
## ② Benchmarks Miss What Judges See

DPO, GRPO and GSPO look alike on benchmarks. Judges tell them apart, and agree on a winner for en-AU and en-UK.

Benchmarks: each dot is one model family and variety on one benchmark (Table 1, explicit thread).



Judges picked the most dialectal of the DPO, GRPO and GSPO outputs from Llama-3.1-8B (Table 4, T4): two native or near-native speakers per variety, and Phi-4 with and without the eWAVE feature list.



## ④ Limits

eWAVE lists the features a variety is known to use; it was not designed to measure how dialectal a text sounds. On en-AU, Phi-4 given the eWAVE feature list picks the dialect-specific output in 81% of judgements, while annotators are split at 48%.

Some features only surface in particular registers, so our generic prompts may not elicit them.

Indian English is consistently the most challenging variety, and has the most active features (38).

What eWAVE features look like, with attestation ratings (A = pervasive, B = common; Appendix B, Table 6)

|       | eWAVE feature             | Example                    | Rating |
|-------|---------------------------|----------------------------|--------|
| en-UK | Possessive me             | He's <b>me</b> brother     | A      |
| en-UK | was/were generalisation   | he <b>were</b> thirsty     | B      |
| en-IN | Progressive with statives | I'm <b>liking</b> this     | A      |
| en-IN | No wh- inversion          | <b>What you are</b> doing? | A      |
| en-AU | Progressive with statives | I'm <b>liking</b> this     | B      |
| en-AU | Sentence-final like       | it was good, <b>like</b>   | B      |



Target varieties: Australian, Indian and Northern British English  
Pretraining: 18 varieties from the International Corpus of English

Variety targeting works, benchmarks do not show it, and a feature-density reward does not make output more dialectal. Closing these gaps needs richer reward designs and more dialectal resources.

All code, checkpoints and dialectal preference datasets are released:

[github.com/surrey-nlp/DiaLLM](https://github.com/surrey-nlp/DiaLLM)

Questions: [j.painter@surrey.ac.uk](mailto:j.painter@surrey.ac.uk)

