Enhancing LLM Translation Performance for Spanish–Valencian through Supervised Fine-Tuning and Reinforcement Learning

Paula Guerrero Castelló
University of the Basque Country (UPV/EHU)
pguerrero005@ikasle.ehu.eus

Abstract

Valencian, the Western Catalan variety used in the Valencian Community of Spain, lacks a dedicated language code in most multilingual machine translation (MT) systems, and is systematically rendered closer to the standard written Eastern Catalan used in Catalonia. We address this gap by adapting TranslateGemma-4B-IT, a 4-billion-parameter instruction-tuned (IT) large language model (LLM) specialized for translation, via three post-training strategies using only public corpora and Quantized Low-Rank Adaptation (QLoRA): (i) supervised fine-tuning (SFT); (ii) Group Relative Policy Optimization (GRPO), a reinforcement learning (RL) technique, with chrF plus a naturalness reward (GRPOv1); and (iii) GRPO with a composite automatic-metric reward (GRPOv2). Our results suggest that reward-function alignment with the target dialect is a key determinant of RL success in low-resource dialectal MT.

1 Introduction

Valencian belongs to the Western branch of the Catalan language family and holds official status alongside Spanish (ES) in the Valencian Community.

Its morphological and lexical properties diverge in important ways from the standard written Eastern Catalan used in Catalonia. Among the most frequent contrasts are verbal morphology (nearly all subjunctive forms differ, e.g. *diga/digui*) and third-person possessives (*seua/seu* vs. *seva/seu*). Some temporal and lexical items also differ: *hui/avui*, *xicotet/petit*, *vore/veure*.

In current NLP practice, however, these two varieties are conflated under the ISO 639-1 code "ca": most parallel corpora assign this label to the standard written Eastern Catalan used in Catalonia, and translation models include no dialect-specific code. Although Valencian does hold a dedicated code in the finer-grained ISO 639-6 standard (vlca), and several open-source software packages already employ this variety, these dialectal identifiers remain unsupported by major MT frameworks.

TranslateGemma-4B-IT (Finkelstein et al., 2026) is a representative example of this limitation. The model conditions on a target_lang_code prompt field, but only "ca" is available for the entire Catalan continuum. Zero-shot prompting with this code predominantly produces Eastern Catalan morphological forms regardless of the user's intended regional target.

Adapting such a model without full fine-tuning, without proprietary data, and without sufficient computational resources poses a realistic challenge for practitioners working on regional varieties, and thus constitutes the central motivation of this paper[1].

We make three specific contributions. First, we establish a reproducible QLoRA-based SFT approach for the ES–VLCA direction, where QLoRA, an efficient fine-tuning method that applies low-rank adapters to quantized models, substantially closes the zero-shot quality gap. Second, we conduct a systematic comparison of two GRPO reward designs for dialectal MT: one that incorporates a learned naturalness classifier, and one relying on automatic reference-based metrics and type-token

[1]Code available at https://github.com/guerreropaula/spanish-valencian-mt-rl.

ratio (TTR). Third, we introduce a *dialectal Valencian score* (DVS) as a complementary evaluation axis alongside standard MT metrics, and use it to demonstrate that the reward design choice has opposing consequences for translation quality and dialect production.

2 Background

Dialectal and low-resource MT. Neural approaches to dialectal variation have progressed from rule-based dialectal preprocessing pipelines (Salloum and Habash, 2012) to unified neural models that utilize multitask learning (Baniata et al., 2018). Nevertheless, much of the research in dialectal MT has primarily focused on Arabic varieties. Regarding the Iberian languages, there is comparatively little prior work, and systems addressing the Catalan–Valencian divergence have received limited attention despite the growing availability of resources from the AINA initiative (Villegas, 2023) and the PlanTL-GOB-ES corpora (Secretaría de Estado de Telecomunicaciones y Sociedad de la Información, 2015). In this context, Apertium (Forcada et al., 2011), an open-source rule-based MT platform currently used by several Valencian institutions, represents one of the main references for Valencian MT.

The most closely related prior work is Galiano Jiménez et al. (2024), who developed systems for the WMT 2024 shared task on translation into low-resource languages of Spain, exploiting Valencian and Catalan as auxiliary transfer-learning languages. Our approach differs in that we target ES–VLCA directly, adapt a translation-specialized LLM via QLoRA-based post-training rather than building dedicated MT pipelines, and introduce a comparison of RL reward designs for dialectal adaptation. The challenge is compounded by the fact that the two varieties share a large common vocabulary, so models trained on standard Eastern Catalan can achieve acceptable metric scores on Valencian test sets while still producing incorrect morphological forms.

RL for MT refinement. Sequence-level training with automatic metrics such as BLEU (Papineni et al., 2002) and ROUGE (Lin and Hovy, 2003) as a reward was proposed by Ranzato et al. (2016) and extended to human preference signals by Kreutzer et al. (2018). Wu et al. (2018) provide a systematic study of RL training strategies for neural MT (NMT), showing that reward shaping—the practice of designing intermediate rewards to guide policy updates—and baseline design are critical for stable training. Work using quality-aware decoding with learning-based metrics such as COMET and BLEURT (Rei et al., 2020; Sellam et al., 2020) has further shown that these metrics capture translation quality better than n-gram overlap alone (Fernandes et al., 2022). In this context, GRPO (Shao et al., 2024), originally developed for mathematical reasoning, eliminates the value network (critic) by normalizing rewards within sampled output groups, therefore reducing memory requirements and making it tractable for billion-parameter models under constrained hardware. Recent work has applied similar RL objectives to GRPO directly to MT, including Feng et al. (2025) and the WMT 2025 submission by Wang et al. (2025), further highlighting the growing adoption of RL-based refinement methods for translation LLMs.

Naturalness in MT. Improving the naturalness and fluency of MT output is a distinct and understudied goal. MT outputs exhibit reduced lexical diversity and increased source-language interference compared to human translation: what is referred to as machine *translationese* (Vanmassenhove et al., 2021). Metrics such as TTR and MTLD are designed to capture such effects. In parallel, Lai et al. (2025) propose an RL-based alignment framework for neural MT that uses COMET and binary classifiers—trained on human translation (HT), machine translation (MT), and original target-language data (OR)—as reward models to approximate human preference. Their findings directly motivate both the TTR component in our GRPOv2 reward and our decision to test a learned HT/MT naturalness classifier as part of the GRPOv1 reward.

Parameter-efficient fine-tuning. Low-Rank Adaptation (LoRA) (Hu et al., 2022) constrains weight updates to low-rank decompositions; combined with 4-bit NF4 quantization as in Quantized LoRA (QLoRA) (Dettmers et al., 2023), it enables adaptation of multi-billion-parameter models on a single GPU, which is the hardware regime assumed throughout this work. We therefore adopt this approach in our experiments.

3 System description

3.1 Training pipeline overview

Figure 1 illustrates the two training pipelines investigated. Both share an SFT stage that adapts the

base model to the ES–VLCA direction; they diverge only in the subsequent GRPO reward design.

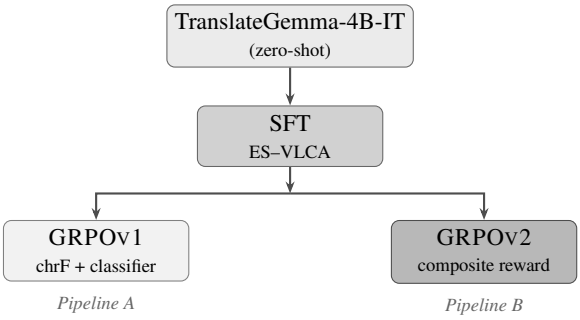

Figure 1: Training pipelines. Both share the SFT stage and differ only in the GRPO reward signal applied thereafter; Pipeline B performs best.

3.2 Base model

TranslateGemma-4B-IT (Finkelstein et al., 2026) is a 4-billion-parameter instruction-tuned model pre-trained to optimize translation quality across 55 languages. Following the authors' specification, we use the structured JSON template with "`ca`" as target language code throughout all experimental conditions; no Valencian-specific code is available.

3.3 Data

SFT corpus. We use the GPLSI AMIC parallel dataset,[2] which provides 50,000 ES–VLCA sentence pairs sourced from news articles published by the *Associació de Mitjans d'Informació i Comunicació* (AMIC). The target Valencian side exposes the model to the morpho-lexical features that distinguish the desired variety from other Catalan varieties.

GRPO corpus. The Spanish source side of the same corpus is reused for RL training: 5,000 sentences for GRPOv1 and 10,000 for GRPOv2. These subsets are randomly sampled from the 50,000 pairs. The smaller corpus for GRPOv1 reflects the higher per-step computational cost of the classifier reward, which made training on more sentences infeasible given the available hardware. Note that this difference in corpus size means the two GRPO variants are not fully matched in compute; readers should bear this in mind when comparing their results directly. Valencian references are used solely to compute reward signals and are never provided to the policy as decoding context.

HT/MT classifier training data. The naturalness classifier, further explained in Section 3.6, is trained on professional human translations drawn from TildeMODEL (general), DOGC (official/legal), and Europarl (parliamentary), paired with MT outputs from OPUS-MT (Tiedemann and Thottingal, 2020) and NLLB-200 (Costa-jussà et al., 2022), yielding 108,000 training and 12,000 validation instances. Each system produces 50% of the translations to prevent system-specific bias. We note that all classifier training data derive from standard Eastern Catalan corpora. This distribution gap between standard Eastern Catalan and Valencian is the central limitation we identify in Section 6. The use of a standard-Catalan classifier was motivated by the absence of sufficiently large Valencian human-translation corpora suitable for classifier training.

Evaluation. All systems are evaluated on the GPLSI ES–VLCA translation test set,[3] a 1,000-sentence held-out set of professional news translations, which ensures adherence to dialectal particularities.

3.4 Supervised fine-tuning

QLoRA targets all attention and feed-forward projection matrices with rank r=16, scaling α=32, and 4-bit NF4 double quantization, yielding 32.8M trainable parameters (0.76% of 4B). The model is trained for three epochs over the 50,000 parallel sentence pairs.

3.5 HT/MT naturalness classifier

We fine-tune RoBERTa-ca (Armengol-Estapé et al., 2021) as a binary classifier estimating $P(\mathrm{HT} \mid \text{hypothesis})$. On the standard-Catalan validation set it attains accuracy 74.7%, F1 74.5%, recall 74.7%, and precision 75.2%. We acknowledge that an F1 of approximately 75% represents a moderate level of reliability; its limitations in the context of reward modeling are discussed in Section 6.

3.6 GRPOv1: hybrid reward

Following SFT, we apply GRPO-based RL to further refine the model. GRPO is a policy-gradient algorithm in which rewards are normalized within a group of sampled outputs for the same input, eliminating the need for a separate value network (Shao

[2] `https://huggingface.co/datasets/gplsi/amic_parallel`

[3] `https://huggingface.co/datasets/gplsi/ES-VA_translation_test`

et al., 2024). For each source sentence, the policy samples $G{=}4$ independent hypotheses. Within-group reward normalization produces advantage estimates used by the GRPO objective with Kullback–Leibler (KL) divergence penalty $\beta{=}0.04$ and learning rate $5{\times}10^{-6}$. In this first variant, the reward combines reference fidelity and estimated naturalness:

$$r = (1-\alpha)\, r_c + \alpha\, r_t \tag{1}$$

where $r_c = \mathrm{chrF}/100$ and $r_t = P(\mathrm{HT}\,|\,\mathrm{hyp})$. To prevent the naturalness term from destabilizing early training, α is held at zero for the first 50 warm-up steps and then increases linearly to 0.3. Training proceeds for 100 steps; the checkpoint at step 80 is retained based on held-out reward.

3.7 GRPOv2: composite automatic-metric reward

GRPOv2 eliminates the classifier dependency, replacing it with a combination of three automatic signals:

$$r = 0.5\,\hat{r}_{\mathrm{chrF}} + 0.3\,\hat{r}_{\mathrm{COMET}} + 0.2\,\mathrm{TTR} - \mathbf{1}[\mathrm{copy}] \tag{2}$$

where $\hat{r}(\cdot)$ denotes min-max normalization to $[0, 1]$. The chrF and COMET components capture surface-level similarity and semantic adequacy, respectively, while the type-token ratio (TTR) promotes lexical diversity and discourages repetitive outputs.

The $\mathbf{1}[\mathrm{copy}]$ term equals 1 when the hypothesis reproduces the source sentence verbatim, enforcing a penalty for copying. This is motivated by the high lexical overlap between Spanish and Valencian as a closely related language pair. Training runs for 200 steps over 10,000 sentences (PPO clip[4] $\varepsilon{=}0.2$, $\beta{=}0.04$). The final model is selected based on the highest mean reward.

4 Evaluation

Standard MT metrics. We report chrF, BLEU, TER, BLEURT, and COMET (Popović, 2015; Papineni et al., 2002; Snover et al., 2006; Sellam et al., 2020; Rei et al., 2020). TER is expressed as a percentage throughout.

[4]PPO (Proximal Policy Optimization) clip is a hyperparameter ε that constrains how far the updated policy may deviate from the previous one in a single gradient step.

System	chrF↑	BLEU↑	TER%↓	BLEURT↑	COMET↑
Baseline	69.02	39.22	40.30	0.258	0.906
SFT	83.16	60.16	22.80	0.524	0.934
GRPOv1	81.65	56.94	23.96	0.481	0.926
GRPOv2	**84.68**	**62.16**	**20.63**	**0.544**	**0.936**

Table 1: Translation quality on the 1,000 ES–VLCA test set. Best in bold. TER: lower is better; reported as percentage.

Dialectal Valencian score (DVS). We introduce the *dialectal Valencian score* (DVS), which measures how frequently a system produces Valencian-specific morpho-lexical surface forms rather than their Eastern Catalan equivalents. We compile a vocabulary of 35 contrastive pairs selected on two criteria: (i) minimum frequency of five occurrences in the 1,000-sentence test set for at least one variant, and (ii) representation across morphological categories (possessives, temporal adverbials, verbal stems, and common lexical alternates), with a roughly equal split across categories. Examples include *seua/seva* (possessive), *hui/avui* (temporal adverb), and *xicotet/petit*, *vore/veure* (lexical alternates). The 35-pair list is provided in Appendix A. DVS is independent of the reference-based metrics above: a system can score high on chrF while still not adhering to Valencian morphology and lexicon.

5 Results

5.1 Automatic translation quality

Table 1 reports metric scores on the test set. SFT delivers the largest single-step absolute improvement: chrF rises 14.1 points, BLEU 20.9 points, and TER falls 17.5 percentage points relative to the zero-shot baseline. GRPOv2 consistently outperforms SFT across all five metrics, with further gains of 1.5 chrF, 2.0 BLEU, and 2.2 TER. By contrast, GRPOv1 regresses below SFT on every metric, a pattern we analyze in Section 6.

5.2 Dialectal Valencian score

Figure 2 and Table 2 present DVS results. SFT achieves the highest overall score (41.0%), confirming that direct supervised exposure to Valencian forms is the most effective lever for dialect adaptation. GRPOv2 retains most of this gain (36.2%), indicating that character-level reward partially reinforces Valencian-specific tokens. GRPOv1, by contrast, collapses to 15.9%, less than half the value of the SFT checkpoint it is initialized from.

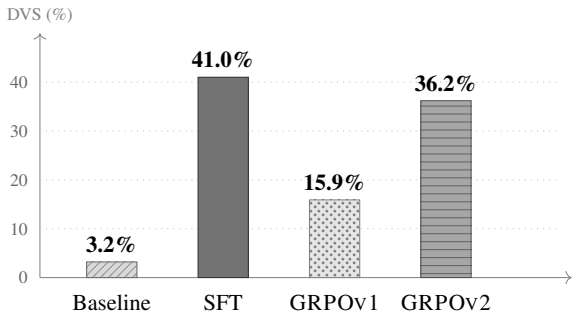

Figure 2: Dialectal Valencian score (DVS): percentage of Valencian forms produced over the 1,000-sentence test set. DVS is computed where at least one contrastive variant is contextually present.

CA	VLCA	Base	SFT	v1	v2
petit	*xicotet*	0%	100%	0%	100%
veure	*vore*	0%	83%	0%	17%
avui	*hui*	0%	71%	0%	86%
seva	*seua*	0%	100%	45%	100%

Table 2: Feature-level DVS for four representative pairs (see Appendix A for the full 35-pair list). GRPOv1 (v1) scores 0% on most Valencian features despite initializing from the SFT checkpoint.

5.3 Qualitative analysis

Table 3 presents a representative example from the test set. All four systems preserve the propositional content; the difference lies in the third-person feminine possessive *seva/seua* ("su"). The baseline and GRPOv1 produce the Eastern Catalan *seva*; SFT and GRPOv2 produce the Valencian form *seua*.

6 Analysis

Classifier-induced dialect suppression in GRPOv1. The performance collapse of GRPOv1 can be traced directly to the distribution mismatch of its classifier. The reward in GRPOv1 penalizes Valencian-specific forms, because the classifier was trained exclusively on standard Eastern Catalan data and consequently assigns lower naturalness scores to Valencian morphology. This behavior is consistent with the finding of Lai et al. (2025) that naturalness classifiers underperform as reward signals when there is a mismatch between the training data and the target-side data used during alignment; we extend this observation to the cross-variety setting. As shown in Table 2, feature-level DVS data support this interpretation: GRPOv1 produces the Valencian possessive *seua* in only 45% of instances despite the SFT model using it in 100%.

System	Output
ES source	*Su participación en este concurso le dio la oportunidad de entrar al mundo artístico.*
Baseline	La **seva** participació … l'oportunitat d'entrar al món artístic.
SFT	La *seua* participació … l'oportunitat d'ingressar al món artístic.
GRPOv1	La **seva** participació … l'oportunitat d'ingressar al món artístic.
GRPOv2	La *seua* participació … l'oportunitat d'ingressar al món artístic.

Table 3: Qualitative comparison. **Bold**: Eastern Catalan form; *italic*: Valencian form.

Composite reward and metric complementarity. The three components of the GRPOv2 reward serve distinct roles. ChrF provides character-level supervision that rewards partial morphological matches, which is particularly important for Valencian-specific suffixes, verb inflections, and possessive forms that differ minimally from their standard Catalan counterparts; this may account for the near-complete preservation of dialectal accuracy relative to the SFT checkpoint. The COMET term introduces semantic evaluation, ensuring that dialectal variation does not come at the cost of meaning preservation. TTR penalizes repetition without imposing any preference over the lexical inventory of the target variety, thus avoiding the classifier-induced dialect suppression documented in GRPOv1. Finally, the copy penalty discourages verbatim source copying, motivated by the high similarity between the language pair. The effectiveness of this configuration suggests that multi-objective RL rewards combining fidelity and semantic quality signals are more robust than single-metric or classifier-based approaches for dialectal low-resource translation.

SFT as the primary dialect-adaptation signal. The highest DVS is achieved by SFT rather than by any RL variant. This reveals a tension in the reward: scalar reward signals provide no explicit incentive for producing Valencian-specific lexical choices. Direct SFT, by contrast, supplies explicit examples of the desired dialectal distribution, thus providing an unambiguous policy gradient toward the desired Valencian variety. Accordingly, future work should prioritize the curation of high-quality, dialect-annotated parallel corpora; enlarging the SFT training data is likely to result in higher returns for dialect adaptation than further RL tuning over resource-constrained settings.

7 Limitations

The training and evaluation data derive from a single domain (news journalism), and performance on legal, administrative, or literary Valencian text remains untested. DVS, while interpretable and useful in this work, captures only lexical surface contrasts and does not account for morphosyntactic divergences (verbal periphrases, clitic ordering, verbal paradigm differences such as subjunctive forms that are equally characteristic of Valencian). Furthermore, the current DVS is primarily based on the written standard established by the Valencian normative authority (Acadèmia Valenciana de la Llengua, 2006); a DVS inclusive of all competing standards would likely yield different results.

The automatic naturalness signals in GRPOv2 (type-token ratio, chrF, COMET) remain shallow proxies for fluency that would ideally be validated against native-speaker judgments. Training hardware constraints bound the total compute invested and may affect reproducibility across configurations. The non-comparability of compute between GRPOv1 and GRPOv2 (5,000 vs. 10,000 training sentences) is an additional caveat when interpreting their relative performance. Finally, our experiments cover a single language pair and domain; the conclusions drawn about reward-function alignment and dialect suppression should therefore be interpreted as initial findings that warrant replication across other dialectal pairs before being further generalized.

8 Conclusion

We have shown that a pre-trained translation-specialized LLM can be adapted to the under-resourced Valencian dialect through lightweight post-training on publicly available corpora and commodity hardware. SFT on 50,000 in-domain pairs constitutes the single largest improvement (+14.1 chrF, +20.9 BLEU) and achieves the highest dialectal form production rate (41.0% DVS). GRPOv2 improves on SFT across all five standard MT metrics and retains 36.2% dialectal adequacy, establishing a strong baseline for the ES–VLCA direction. The regression of GRPOv1 has implications beyond this language pair: a naturalness classifier trained on a related standard variety is actively harmful when applied as an RL reward signal for dialectal adaptation, since, as demonstrated, it systematically penalizes the target dialect's characteristic forms. This outcome underscores the importance of aligning reward models with the actual target variety. Future work should address this by (i) training the naturalness classifier exclusively on high-quality professional Valencian translations; (ii) scaling the SFT corpus via controlled data augmentation; (iii) extending DVS to morphosyntactic features beyond surface contrasts, while accounting for competing Valencian written norms; and (iv) conducting human evaluation with native Valencian speakers to disentangle automatic-metric artifacts from real translation quality improvements.

Acknowledgments

We sincerely thank Antonio Toral, Nora Aranberri, Mikel Forcada, Gorka Azkune, and Eneko Agirre for their valuable support and guidance throughout the development of this work.

References

Acadèmia Valenciana de la Llengua. 2006. *Gramàtica normativa valenciana*. Acadèmia Valenciana de la Llengua, València.

Armengol-Estapé, Jordi, Casimiro Pio Carrino, Carlos Rodriguez-Penagos, Ona de Gibert Bonet, Carme Armentano-Oller, Adriana Gonzalez-Agirre, Maite Melero, and Marta Villegas. 2021. Are Multilingual Models the Best Choice for Moderately Underresourced Languages? A Comprehensive Assessment for Catalan. In *Findings of ACL-IJCNLP 2021*, pages 4933–4946.

Baniata, Laith H., Seyoung Park, and Seong-Bae Park. 2018. A Neural Machine Translation Model for Arabic Dialects that Utilises Multitask Learning (MTL). *Computational Intelligence and Neuroscience*, 2018.

Costa-jussà, Marta R., Angela Fan, Shruti Bhosale, et al. 2022. No Language Left Behind: Scaling Human-Centered Machine Translation. *arXiv preprint arXiv:2207.04672*.

Dettmers, Tim, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. QLoRA: Efficient finetuning of quantized LLMs. *Advances in Neural Information Processing Systems*, 36.

Feng, Zhaopeng, Shaosheng Cao, Jiahan Ren, Jiayuan Su, Ruizhe Chen, Yan Zhang, Zhe Xu, Yao Hu, Jian Wu, and Zuozhu Liu. 2025. MT-R1-Zero: Advancing LLM-based Machine Translation via R1-Zero-like Reinforcement Learning. In *Findings of ACL: EMNLP 2025*, pages 18685–18702, Suzhou, China. Association for Computational Linguistics.

Fernandes, Patrick, António Farinhas, Ricardo Rei, José G. C. de Souza, Perez Ogayo, Graham Neubig, and André F. T. Martins. 2022. Quality-Aware Decoding for Neural Machine Translation. In *Proceedings of*

NAACL 2022, pages 1396–1412, Seattle. Association for Computational Linguistics.

Finkelstein, Mara, Isaac Caswell, Tobias Domhan, Jan-Thorsten Peter, Juraj Juraska, Parker Riley, Daniel Deutsch, Geza Kovacs, Cole Dilanni, Colin Cherry, Eleftheria Briakou, Elizabeth Nielsen, Jiaming Luo, Kat Black, Ryan Mullins, Sweta Agrawal, Wenda Xu, Erin Kats, Stephane Jaskiewicz, Markus Freitag, and David Vilar. 2026. TranslateGemma Technical Report. *arXiv preprint arXiv:2601.09012*.

Forcada, Mikel L., Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, and Francis M. Tyers. 2011. Apertium: a free/open-source platform for rule-based machine translation. *Machine Translation*, 25(2):127–144.

Galiano Jiménez, Aaron, Víctor M. Sánchez-Cartagena, Juan Antonio Pérez-Ortiz, and Felipe Sánchez-Martínez. 2024. Universitat d'Alacant's Submission to the WMT 2024 Shared Task on Translation into Low-Resource Languages of Spain. In *Proceedings of the Ninth Conference on Machine Translation*, pages 885–891, Miami, Florida, USA. Association for Computational Linguistics.

Hu, Edward J., Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, and Weizhu Chen. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *Proceedings of the International Conference on Learning Representations*.

Kreutzer, Julia, Shahram Khadivi, Evgeny Matusov, and Stefan Riezler. 2018. Can Neural Machine Translation be Improved with User Feedback? In Bangalore, Srinivas, Jennifer Chu-Carroll, and Yunyao Li, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 3 (Industry Papers)*, pages 92–105, New Orleans - Louisiana, June. Association for Computational Linguistics.

Lai, Huiyuan, Esther Ploeger, Rik Van Noord, and Antonio Toral. 2025. Multi-perspective Alignment for Increasing Naturalness in Neural Machine Translation. In Che, Wanxiang, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 28071–28084, Vienna, Austria, July. Association for Computational Linguistics.

Lin, Chin-Yew and Eduard Hovy. 2003. Automatic evaluation of summaries using N-gram co-occurrence statistics. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, NAACL '03, pages 71–78, Edmonton, Canada. Association for Computational Linguistics.

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: a Method for Automatic Evaluation of Machine Translation. In Isabelle, Pierre, Eugene Charniak, and Dekang Lin, editors, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA, July. Association for Computational Linguistics.

Popović, Maja. 2015. chrF: character N-gram F-score for automatic MT evaluation. In Bojar, Ondřej, Rajan Chatterjee, Christian Federmann, Barry Haddow, Chris Hokamp, Matthias Huck, Varvara Logacheva, and Pavel Pecina, editors, *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pages 392–395, Lisbon, Portugal, September. Association for Computational Linguistics.

Ranzato, Marc'Aurelio, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2016. Sequence Level Training with Recurrent Neural Networks. In *Proceedings of the International Conference on Learning Representations*.

Rei, Ricardo, Craig Stewart, Ana C Farinha, and Alon Lavie. 2020. COMET: A Neural Framework for MT Evaluation. In Webber, Bonnie, Trevor Cohn, Yulan He, and Yang Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2685–2702, Online, November. Association for Computational Linguistics.

Salloum, Wael and Nizar Habash. 2012. Elissa: A Dialectal to Standard Arabic Machine Translation System. In Kay, Martin and Christian Boitet, editors, *Proceedings of COLING 2012: Demonstration Papers*, pages 385–392, Mumbai, India, December. The COLING 2012 Organizing Committee.

Secretaría de Estado de Telecomunicaciones y Sociedad de la Información. 2015. Plan de Impulso de las Tecnologías del Lenguaje. Technical report, Ministerio de Industria, Energía y Turismo.

Sellam, Thibault, Dipanjan Das, and Ankur Parikh. 2020. BLEURT: Learning Robust Metrics for Text Generation. In Jurafsky, Dan, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7881–7892, Online, July. Association for Computational Linguistics.

Shao, Zhihong, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models. *arXiv preprint arXiv:2402.03300*.

Snover, Matthew, Bonnie Dorr, Richard Schwartz, Linnea Micciulla, and John Makhoul. 2006. A Study of Translation Edit Rate with Targeted Human Annotation. In *Proceedings of AMTA 2006*, pages 223–231.

Tiedemann, Jörg and Santhosh Thottingal. 2020. OPUS-MT – building open translation services for the world. In Martins, André, Helena Moniz, Sara Fumega, Bruno Martins, Fernando Batista, Luisa Coheur, Carla Parra, Isabel Trancoso, Marco Turchi, Arianna Bisazza, Joss Moorkens, Ana Guerberof, Mary Nurminen, Lena Marg, and Mikel L. Forcada, editors, *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, pages 479–480, Lisboa, Portugal, November. European Association for Machine Translation.

Vanmassenhove, Eva, Dimitar Shterionov, and Matthew Gwilliam. 2021. Machine Translationese: Effects of Algorithmic Bias on Linguistic Complexity in Machine Translation. In Merlo, Paola, Jorg Tiedemann, and Reut Tsarfaty, editors, *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2203–2213, Online, April. Association for Computational Linguistics.

Villegas, Marta. 2023. El projecte AINA, la IA i les tecnologies del llenguatge. *Terminàlia*, 27:80–84.

Wang, Hao, Linlong Xu, Heng Liu, Yangyang Liu, Xiaohu Zhao, Bo Zeng, Longyue Wang, Weihua Luo, and Kaifu Zhang. 2025. Marco Large Translation Model at WMT2025: Transforming Translation Capability in LLMs via Quality-Aware Training and Decoding. In *Proceedings of WMT 2025*.

Wu, Lijun, Fei Tian, Tao Qin, Jianhuang Lai, and Tie-Yan Liu. 2018. A Study of Reinforcement Learning for Neural Machine Translation. In *Proceedings of EMNLP 2018*, pages 3612–3621, Brussels, Belgium. Association for Computational Linguistics.

A DVS Contrastive Vocabulary

Table 4 lists all 35 contrastive pairs used to compute the *dialectal Valencian score*. The pairs are grouped by morphological category. All pairs meet the minimum frequency threshold of five occurrences in the 1,000-sentence test set. The list reflects primarily the AVL written standard.

Category	CA form	VLCA form
Determiners & possessives		
	aquesta	*esta*
	aquest	*este*
	aquestes	*estes*
	aquests	*estos*
	seva	*seua*
	seves	*seues*
	darrer	*últim*
	darrers	*últims*
	darrera	*última*
Verbal stems & inchoative endings		
	tenir	*tindre*
	obtenir	*obtindre*
	veure	*vore*
	segueix	*seguix*
	segueixen	*seguixen*
	requereix	*requerix*
	divideix	*dividix*
	constitueixen	*constituïxen*
	absorbeixen	*absorbixen*
Lexical alternates		
	nens	*xiquets*
	nen	*xiquet*
	nena	*xiqueta*
	nenes	*xiquetes*
	petit	*xicotet*
	petits	*xicotets*
	petita	*xicoteta*
	feina	*faena*
	feines	*faenes*
	cop	*colp*
	cops	*colps*
	avui	*hui*
	servei	*servici*
	serveis	*servicis*
	mirall	*espill*
	tomàquet	*tomaca*
	tomàquets	*tomaques*

Table 4: Full DVS contrastive vocabulary grouped by morphological category. Pairs reflect the AVL written standard. Each entry represents a CA–VLCA surface-form pair; DVS is computed over sentences where at least one variant of each pair is contextually present.