# A Comparative Study of Multilingual Fine-tuning and Prompting for Automatic Text Readability Classification in Galician

**Sandra Rodríguez Rey, Marcos Garcia**
Centro Singular de Investigación en Tecnoloxías Intelixentes, Universidade de Santiago de Compostela
Rúa Jenaro de la Fuente Domínguez, 15782, Santiago de Compostela
{sandrarodriguez.rey,marcos.garcia.gonzalez}@usc.gal

**Abstract**

Despite advancements in automatic readability assessment, low-resource languages such as Galician remain under-explored. This study addresses this gap by presenting a comparative study of readability assessment techniques in Galician, including fine-tuning of encoder models as well as prompting strategies using large generative models. Due to the scarcity of native Galician resources, neural machine translation was employed to generate synthetic Galician data. The analysis begins with BERT-based monolingual models trained on the synthetic data. For multilingual models, the impact of using original versus translated data was compared in order to assess the effects of translation-based augmentation. Finally, several LLMs were evaluated using zero-shot and few-shot prompting methods. The results indicate that generative models are not yet competitive with encoder models tuned for text classification in Galician, and that data generated through machine translation improves the performance of monolingual models but has little effect on multilingual models.

**Keywords:** Automatic Readability Assessment, Galician, Data Augmentation, Machine Translation, Text Classification.

## 1. Introduction

Various strategies are currently employed in the development of automatic text classifiers, from linguistically-informed machine learning models to deep learning, hybrid approaches, and strategies involving Large Language Models (LLMs). Recent work indicates that model adjustment using fine-tuning techniques remains one of the most widely used and recommended approaches for multi-class classification tasks (Trokhymovych et al., 2024; Imperial et al., 2025). Concurrently, recent advances in generative artificial intelligence have inspired the exploration of using large language models for classification tasks through instructions and examples (Pungeršek et al., 2025; Kostina et al., 2025).

For similar tasks, and particularly for languages with limited data, some studies suggest that certain data augmentation strategies can improve results (Ziyaden et al., 2024). However, other studies, especially those involving automatically translated data, reveal that the results obtained are slightly inferior to those obtained with original texts (Rehan et al., 2023). Nevertheless, the impact of using translated data may vary depending on the type of model used, for example, monolingual versus multilingual, therefore motivating a focused analysis of this strategy.

In this context, this paper presents a comparative study of strategies to develop automatic text classifiers by readability level in Galician, including original data for this language as well as data augmentation techniques. To this end, this study uses datasets that include texts categorized into four distinct readability levels. These datasets cover multiple textual genres and are specifically designed for developing and evaluating automatic classifiers. The study focuses on three areas: the use of fine-tuned Transformers models, both monolingual and multilingual; the analysis of the impact of using data from other languages, both in their original version and automatically translated into Galician; and the evaluation of large language models provided with instructions and examples.

Thus, the main contributions of this work are as follows: (i) A comparative study of strategies for developing Galician automatic text classifiers by readability level, including fine-tuned and large language models, (ii) an analysis of the impact of data augmentation strategies using corpora from other languages in their original form and in an automatically translated version in Galician in both monolingual and multilingual models, and (iii) the publication of the best models in open repositories, such as Hugging Face[1].

## 2. Related Work

Early work on readability assessment relied on formula-based metrics grounded in surface features such as sentence length and word complexity (Flesch, 1948; Harry and Laughlin, 1969; Dale and Chall, 1948), followed by mixed feature approaches leveraging tools like Coh-Metrix to capture richer linguistic indicators (Crossley et al., 2007), and more

[1] `https://huggingface.co/sandrarrey`



*Proceedings of the Joint Workshop on Readability and Text Simplification (READIxTSAR) @ LREC 2026*, pages 101–120
11 May 2026. 

recently by supervised machine learning and deep learning architectures for automatic assessment (Dell'Orletta et al., 2014; Madrazo Azpiazu and Pera, 2019).

Several recent studies on classifying texts according to their readability at multiple levels indicate that fine-tuning pre-trained models remains one of the most widespread and effective approaches (Trokhymovych et al., 2024; Imperial et al., 2025; Ribeiro et al., 2024; Coban et al., 2024; Vásquez-Rodríguez et al., 2022; Hernandez et al., 2022). In practice, this approach constitutes the dominant method for automatically classifying texts by readability level. In parallel, recent advances in generative artificial intelligence have motivated the exploration of using large language models for classification tasks with instructions and examples in the prompt (Pungeršek et al., 2025; Kostina et al., 2025). However, various studies conclude that fine-tuned models offer clearly superior performance to generative models in multi-class classification tasks (Imperial et al., 2025; Pungeršek et al., 2025). Other studies qualify this conclusion, pointing out that the results obtained with generative models can be competitive, although often at the expense of higher computation times (Kostina et al., 2025).

Using corpora annotated by readability levels is essential for training and evaluating these systems. Related work has collected reference corpora and datasets in various languages (Schwarm and Ostendorf, 2005; Wilkens et al., 2022; Ribeiro et al., 2024; Quispesaravia et al., 2016). The iRead4Skills dataset is a recent resource that stands out. It is a multilingual dataset with texts classified by complexity levels in Spanish, Portuguese, and French (Pintard et al., 2024). For languages with limited resources, some studies have been designed to facilitate text evaluation in languages with less available annotated data, such as Basque (Gonzalez-Dios et al., 2018) or Galician (Rodríguez Rey and Garcia, 2025).

In contexts with limited corpus availability, data augmentation strategies are a fundamental resource for improving classifier performance. One of the most widely studied techniques is reusing datasets from other languages, either unchanged or adapted through machine translation. This is done both to obtain readability corpora for low-resource languages (Sibeko, 2024) and to enhance text classification tasks in general (Ziyaden et al., 2024). While some studies show that models trained with machine translated augmented data can significantly outperform those trained with only original data (Ziyaden et al., 2024), other studies conclude that translated texts generate slightly lower performance in models compared to original texts (Rehan et al., 2023). This apparent contradiction in results shows that the effectiveness of this strategy depends on the type of model used, the quality of the translation, and the specific language. Therefore, it is necessary to investigate how combining data augmentation strategies with monolingual or multilingual approaches, such as mBERT or XLM-RoBERTa, affects the performance of classifiers in languages with few resources (Pakray et al., 2025).

Automatic readability assessment plays a vital role in language education and digital accessibility (Vajjala, 2022). However, despite its recognized significance for learning and inclusion, notable gaps remain in the research on text readability in languages with limited resources. While the broader field has seen recent progress, strategies for developing high-quality automatic classifiers for these languages have hardly been analyzed. For Galician, building robust models requires a sufficient amount of annotated data. However, only a small corpus is available for this task (Rodríguez Rey and Garcia, 2025). This makes reusing multilingual datasets and exploring data augmentation techniques necessary. This is where this work comes in.

This study aims to advance the development of automatic text classifiers for Galician by comparing different models and analyzing the impact of using translated versus original texts in monolingual and multilingual scenarios. Beyond a technical perspective, this work aims to contribute to the design of specialized pedagogical tools and to foster the creation and adaptation of educational materials for Galician language teaching. One practical application of these classifiers, for example, is that they could help teachers and independent learners easily find and select reading materials that are just right for their level. This is particularly important for languages that have recently been incorporated into formal education, since even native adult speakers may have difficulty achieving proficiency and developing consistent reading habits in Galician.

## 3. Research Questions (RQs) and Hypotheses (Hs)

Building on prior studies which show that fine-tuning encoder models is effective for multi-class text classification (Trokhymovych et al., 2024; Pungeršek et al., 2025; Imperial et al., 2025), and that multilingual models and data augmentation improve performance (Ziyaden et al., 2024), we formulate the following research questions and corresponding hypotheses:

- RQ1: Does translation-based data augmentation improve the performance of fine-tuned models?

H1: Monolingual models would benefit most from the introduction of data in the target language. Multilingual models could perform better with a larger volume of texts translated into the same language in which they will be evaluated. However, if the languages of the original resources are included in the multilingual model, translation may not be necessary.

- RQ2: Considering the recent advancements in generative models, is it worthwhile to fine-tune a model for multi-class text classification, or can generative models achieve competitive or even superior performance to fine-tuned models?
  H2: As indicated by related work (Imperial et al., 2025), model fine-tuning remains the most effective technique for multi-class classification tasks. However, recent advances in generative AI, especially instructed models, could compete with this technique (Pungeršek et al., 2025; Kostina et al., 2025), as these models can work with instructions and few examples, saving work.

## 4. Resources: datasets and models

The following datasets and language models were used to conduct experiments comparing the performance of different tools for classifying texts according to their readability in Galician.

### 4.1. Datasets

The following corpora were utilized. A proprietary script was used to convert them into CSV datasets compatible with the libraries used for model development.

**Corlega corpus**[2] : This corpus includes 480 Galician texts (145,854 tokens) from various subgenres and themes, such as social media, fiction and nonfiction literature, professional websites, ads, political discourse, and legal texts, extracted from diverse online sources. The texts are aimed at adult Galician speakers looking to improve their skills and learners of Galician. One expert annotator classified the documents into four levels (first to fourth). These levels were defined by a set of linguistic descriptors describing the lexical-conceptual, verbal, syntactic, cohesion, and textual characteristics of each level. The levels are based on those defined in the iRead4Skills multilingual dataset (see below) and have been adapted primarily according to the specifications of the Celga (the standard system for certifying proficiency in Galician) and CEFR (Common European Framework of Reference for Languages) levels (Rodríguez Rey and Garcia, 2025).

**iRead4Skills Dataset 1:** This dataset includes three corpora of Spanish, Portuguese, and French texts classified by complexity (Pintard et al., 2024). These corpora consist of 2,000 to 3,000 texts per language from various subgenres and themes intended for native adult speakers. The texts are classified into four levels (very easy, easy, clear, and more complex), three of which are defined by experts (levels one to three), and a fourth reference level that includes texts that are more complex than the previous ones (Monteiro et al., 2024).

**Galician translation of the iRead4Skills Dataset 1:** The dataset previously mentioned was automatically translated into Galician using SalamandraTA[3], and this translated version was also used in some experiments.

#### 4.1.1. Datasets splitting

The Corlega corpus was used for the evaluation, and the iRead4Skills Dataset 1 and its Galician translation were used for training. The Corlega corpus was divided into two parts: validation (25%) and evaluation (75%). Using a proprietary script, the texts were chosen randomly to ensure a balance in the number of texts per level in each part. Given the imbalance in the number of texts per level in the training sets, the final amount was adjusted using an oversampling strategy. Texts were randomly duplicated from levels with fewer documents to equalize the model's recognition capacity across all levels. An automated consistency check was applied to the translated dataset, and documents were excluded when the length ratio of the original and translated texts differed by more than 30%. In total, 119 translated texts were deleted, accounting for 1.56% of the texts. The final size of the datasets is shown in Table 1.

### 4.2. Models

The following models, both encoders and LLMs, were used:

#### 4.2.1. Encoders

The following monolingual and multilingual models, extracted from the Transformers library in Hugging

[2]The Corlega corpus is available under a CC BY-NC-ND 4.0 license for research and reproducibility purposes at Zenodo (Rodríguez Rey and Garcia, 2025). More information about corpus creation is available at Rodríguez Rey and Garcia (2025).

[3]https://huggingface.co/BSC-LT/salamandraTA-7b-instruct

| Language | split | O.T. | T.T. |
|---|---|---|---|
| GL | Test | 360 | - |
| GL | Valid | 120 | - |
| PT+SP+FR | Train | 7677 | 9728 |
| GL(MT)* | Train | 7558 | 9556 |

Table 1: Total number of texts in the original datasets *(O.T.)* and total amount of texts with oversampling *(T.T.)*. *GL(MT) refers to the PT+SP+FR dataset translated into Galician.

Face (Wolf et al., 2020), were used as a starting point to design a baseline model and perform the fine-tuning experiments:

- BERT-gl (Garcia, 2021), using its two variants: *small* and *base*.
- Bertinho (Vilares et al., 2021), also in its *small* and *base* variants.
- mBERT (Devlin et al., 2019), a multilingual BERT-base model.
- XLM-RoBERTa (Conneau et al., 2019) in its *base* and *large* variants.

#### 4.2.2. Generative models

To explore the potential of LLMs for the proposed classification task, the following models were used:

- Llama-3.1-8B-Instruct[4] (Grattafiori et al., 2024). It includes closely related Romance languages, such as Spanish, Portuguese, French, and Italian, but not Galician.
- Llama3.1-Carvalho-PT-GL 8B[5], a continuation of the training of Llama 3.1 8B (Grattafiori et al., 2024) on a large monolingual Galician corpus (de Dios-Flores et al., 2024).
- Gemma-3-4b-it[6] (Gemma Team, 2025), a multilingual model that includes 140 languages and has been trained to perform various text and image generation tasks. Although the list of supported languages is not explicitly stated, it can be assumed that Galician is included because the model was evaluated using the Flores-101 benchmark (Goyal et al., 2022), which includes Galician.
- Qwen2.5-7B-Instruct[7] (Yang et al., 2024; Team, 2024), the best-performing model for Galician in the IberBench Leaderboard (Ángel González et al., 2026)[8].
- GPT-5[9], as an example of state-of-the-art proprietary model, used via ChatGPT[10].

## 5. Experiments and results

This section introduces a set of baseline experiments followed by the results of the fine-tuned models as well as of the LLMs. The experiments were performed on a standard server equipped with two NVIDIA Hopper H100 GPUs, each with 80 GB of memory, and 384 GB of RAM.

### 5.1. Baseline

As a baseline for comparison with all other models, a simple hybrid approach which combines surface features and information from Transformer models to train Random Forest classifiers was implemented. As textual features, the mean, median, mininum and maximum sentence and word length of each document were included, totaling 8 features. From Transformers, a document vector, and surprisal were included. The former is constructed by averaging all sentence vectors in the document, each sentence being represented as the mean of all its sub-tokens (Imperial, 2021). Regarding surprisal, the method proposed by Kauf and Ivanova (2023) for bidirectional models was followed, using the implementation provided by `minicons`[11] (Misra, 2022).

The classifiers were trained with the Random Forest implementation of `scikit-learn`[12], using the datasets presented in Section 4.1 and the encoder models mentioned in Section 4.2.1. During training, different estimators sizes (50 to 600) were evaluated and the best configuration based on the performance on the validation set was selected.

The results of the experiments with the baseline models are shown in Table 2. For these baseline models, accuracy for performance comparison is reported.

### 5.2. Fine-tuning encoder models

To answer RQ1, three types of experiments were conducted to fine-tune the Transformer models

[4] https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct
[5] https://huggingface.co/Nos-PT/Llama-Carvalho-PT-GL
[6] https://huggingface.co/google/gemma-3-4b-it
[7] https://huggingface.co/Qwen/Qwen2.5-7B-Instruct
[8] https://huggingface.co/spaces/iberbench/leaderboard
[9] https://openai.com/es-ES/index/introducing-gpt-5/
[10] https://chatgpt.com/
[11] https://github.com/kanishkamisra/minicons
[12] https://scikit-learn.org

| Model | Dataset | Accuracy |
|---|---|---|
| BERT-small | GL(MT) | **0.51** |
| BERT-base | GL(MT) | 0.49 |
| Bertinho-base | GL(MT) | 0.47 |
| Bertinho-small | GL(MT) | 0.44 |
| XLM-base | PT+SP+FR | *0.50* |
| mBERT-base | PT+SP+FR | 0.48 |
| XLM-large | PT+SP+FR | 0.47 |
| XLM-base | GL(MT) | *0.48* |
| mBERT-base | GL(MT) | *0.48* |
| XLM-large | GL(MT) | 0.46 |

Table 2: Best results from experiments with baseline models trained with authentic multilingual data (PT+SP+FR) and translated data (GL(MT)), evaluated on the test set. The best results are shown in bold. In italics, the best results in each scenario.

mentioned in Section 4.2.1. The models and training datasets were divided into monolingual and multilingual sets for the different experiments.

**Monolingual:** These experiments were conducted using the monolingual models in Galician, BERT-gl and Bertinho, fine-tuned using the Galician-translated dataset.

**Multilingual with authentic data:** In this case, the experiments were carried out using the multilingual encoders (mBERT and XLM-RoBERTa). Adjustments were made using the original dataset, which contained classified texts in Portuguese, Spanish, and French.

**Multilingual with translated data:** Here the same multilingual models adjusted using the Galician-translated dataset were used.

A standard fine-tuning methodology was employed, consisting in adjusting the model parameters for text classification based on a training corpus. Several hyperparameters related to batch size (8, 16, and 32) and five learning rate values (from 1e-5 to 5e-5) were explored. The models were adjusted with different combinations of these hyperparameters for five epochs and then evaluated using the validation set. The best configurations were selected for each model based on accuracy, which were then evaluated using the test set[13].

Table 3 shows the results of the fine-tuning experiments. The best results for each model were chosen based on accuracy.

[13]Since the results on the validation set were not conclusive for some models (for example, there was a tie for several configurations of the same model), more than one configuration for some models were also evaluated on the test set.

### 5.3. Zero-shot and few-shot LLM prompting

Various experiments were carried out to compare the performance of generative models' potential for this task, employing a standardized methodology and focusing on answering RQ2. In the following, we first present the generic instruction included in the prompts, and then introduce the zero- and few-shot prompting methods.

**Instructions:** The prompts were developed following OpenAI's best practices for prompt engineering[14] and were written in Spanish[15]. Each prompt included role, objective, classification level characteristics, step-by-step task instructions, context, expected model inputs, output format and criteria, restrictions, and common errors to avoid. All models were asked to classify input texts into the four described levels. Each prompt specified:

- The description of the four levels, including their lexical-conceptual, syntactic, verbal, and cohesion features, as well as the most frequent text genres associated with each level[16].
- The requirement to analyze the provided examples (when available).
- The task of classifying the texts from a CSV file by assigning a numerical level from 1 to 4.
- The input data: level descriptions, texts to classify, and, in some cases, example texts.
- The expected output: a CSV table containing the predicted level for each text.
- Restrictions and errors to avoid: inventing characteristics and levels, randomly assigning a level, and responding with information that is not strictly a number from one to four.

**Zero-shot strategy:** In the experiments, only a prompt containing the specified instructions was used, and a set of test texts was provided for level prediction.

[14]https://tinyurl.com/openai-best-practices-prompt

[15]Since LLMs tend to perform better in high-resource languages (Xuan et al., 2025), we conducted these experiments in Spanish, which is significantly more representative than Galician and co-exists with it in the same territory.

[16]This information was extracted directly from the appendix published in Rodríguez Rey and Garcia (2025).

| Model | Dataset | Acc | NAcc | Prec | Rec | F1 |
|---|---|---|---|---|---|---|
| Bertinho-base | GL(MT) | **0.55** | **0.95** | *0.56* | **0.55** | **0.54** |
| BERT-base | GL(MT) | 0.54 | **0.95** | *0.56* | 0.49 | 0.50 |
| Bertinho-small | GL(MT) | 0.52 | 0.94 | 0.52 | 0.51 | 0.52 |
| BERT-small | GL(MT) | 0.51 | 0.94 | 0.51 | 0.51 | 0.51 |
| XLM-base | GL(MT) | *0.54* | 0.92 | *0.54* | *0.53* | *0.53* |
| XLM-large | GL(MT) | 0.52 | 0.92 | 0.51 | *0.53* | 0.51 |
| mBERT | GL(MT) | 0.51 | *0.94* | 0.52 | 0.47 | 0.48 |
| XLM-base | PT+SP+FR | 0.53 | *0.93* | 0.56 | *0.50* | *0.52* |
| XLM-large | PT+SP+FR | *0.54* | 0.91 | **0.58** | *0.50* | 0.50 |
| mBERT | PT+SP+FR | 0.50 | 0.91 | 0.52 | 0.46 | 0.47 |

Table 3: Best results from fine-tuning monolingual models with translated data (GL(MT)), multilingual models with authentic data (PT+SP+FR), and multilingual models with translated data (GL(MT)). *Acc* refers to accuracy, *NAcc* to neighbor or adjacent accuracy, *Prec* to precision, *Rec* to recall, and *F1* to the F-score obtained in the test set. The best results are shown in bold. In italics, the best results in each scenario.

**Few-shot strategy:** In this scenario, 13 examples of classified texts were added to the general instructions within the same prompt. The texts were distributed by level as follows: five examples of level 1, four examples of level 2, three examples of level 3, and one example of level 4. The limited number of examples is due to the maximum size allowed for the prompt. The total size of the examples at each level (one to three) is similar. Only one level 4 example was included because this level is not defined by experts; it is simply a reference for texts that are more complex than those of the defined levels.

Table 4 shows the results of the experiments with generative models.

## 6. Discussion of results and comparative study

Some ideas and trends can be drawn from the results obtained in the different experiments. These results allow us to answer the research questions and to validate or reject the established hypotheses. All calculated metrics were considered in the evaluation, with accuracy serving as the primary reference for selecting the best-performing models.

The baseline models, a hybrid model combining surface features and information extracted from encoder models, achieved similar scores with both original and translated data, as well as with both monolingual and multilingual models. The best results were 0.51 accuracy with a monolingual model and translated data and 0.50 accuracy with a multilingual model and original data. The results suggest that multilingual models perform slightly better with original data, but better results were obtained in only half of the models and the difference was minimal. Additionaly, there are no clear trends regarding the influence of model size on results.

Experiments with fine-tuned models reveal similar trends to those observed with baseline models. There is little difference in the results of monolingual and multilingual models or between models fine-tuned with original or translated data. The best results across all metrics were achieved with a Bertinho base model fine-tuned with translated data. This was followed by a BERT-base model, that had nearly identical precision results but lower recall. This implies that the F1 score is also lower than that of the Bertinho-base model. For multilingual models, the results for mBERT and XLM-base with original and translated data are very similar, though the translated models have a slight advantage. However, XLM-large with original data has better precision (the highest: 0.58), although lower recall than XLM-base with translated data, which is generally the best multilingual model.

Experiments with generative models developed in this study did not reach the minimum values established by baseline models. This seems to indicate that these models, with a standard methodology, are not currently the best option for multi-class text classification tasks. However, more advanced strategies (Liu et al., 2024; Yousefiramandi and Cooney, 2025) can be explored in the future. As expected, the results show that few-shot technique improves the performance of almost all models, except for Llama3.1 Instruct in terms of accuracy and neighbor accuracy, and Gemma-3 in terms of precision. The best model in terms of accuracy with the zero-shot technique is Llama3.1 Instruct, though Gemma-3 is the best model in general terms. Gemma-3 achieves an accuracy and adjacent accuracy slightly below that of Llama, but significantly outperforms it in the rest of the metrics. As anticipated from prior information regarding its high performance in classification tasks, the best model with the few-shot technique is Qwen.

To gain a deeper understanding of how the model

| | Zero-shot | | | | | Few-shot | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Acc** | **NAcc** | **Prec** | **Rec** | **F1** | **Acc** | **NAcc** | **Prec** | **Rec** | **F1** |
| Llama3.1 Instruct | *0.37* | *0.86* | 0.29 | 0.31 | 0.28 | 0.35 | 0.72 | 0.32 | 0.44 | 0.27 |
| Carvalho-PT-GL | 0.31 | 0.82 | 0.24 | 0.27 | 0.23 | 0.37 | 0.83 | **0.55** | 0.32 | 0.28 |
| Gemma-3 | 0.36 | 0.82 | *0.42* | *0.37* | *0.34* | 0.40 | 0.84 | 0.38 | 0.41 | 0.35 |
| Qwen2.5 Instruct | 0.30 | 0.74 | 0.37 | 0.32 | 0.27 | **0.43** | **0.88** | 0.40 | **0.45** | **0.41** |
| GPT-5 | 0.16 | 0.58 | 0.36 | 0.26 | 0.13 | 0.30 | 0.78 | 0.28 | 0.26 | 0.26 |

Table 4: Best results from experiments with generative models evaluated on the test set. *Acc* refers to accuracy, *NAcc* to neighbor or adjacent accuracy, *Prec* to precision, *Rec* to recall, and *F1* to the F-score obtained in the test set. The best results are shown in bold. In italics, the best results in each scenario.

behaves at different levels, a confusion matrix for the best-performing model in each fine-tuning experimental scenario was created. For this analysis, the Bert-base model adjusted with translated data, the XLM-base model adjusted with translated data, and the XLM-large model with data in its original language were selected. Information from the models' predictions on the test was used to create the confusion matrices in Figure 1. The conclusion that Bertinho-base is the best-performing model is reinforced by these visualizations, which complement the quantitative metrics. It classified very few texts with a difference of more than one level and showed better prediction of level 4, which had fewer original texts in its adjustment. Of the two multilingual models selected, XLM-base with translated data made the best predictions. XLM-large, when trained on the original dataset, demonstrates a pronounced error concentration in levels 3 and 4.

Table 5 shows in greater detail the differences in performance by level between the two best models, Bertinho-base and XLM-base, which were both fine-tuned with translated data. Bertinho-base outperforms XLM-base on all metrics and levels, with the exception of level 2 predictions, in which XLM-base surpasses Bertinho-base.

| **L** | **Bertinho-base** | | | **XLM-base** | | |
|---|---|---|---|---|---|---|
| | **Prec** | **Rec** | **F1** | **Prec** | **Rec** | **F1** |
| 1 | 0.70 | 0.67 | **0.68** | 0.65 | 0.62 | 0.63 |
| 2 | 0.45 | 0.53 | 0.49 | 0.46 | 0.59 | **0.52** |
| 3 | 0.56 | 0.46 | **0.50** | 0.55 | 0.44 | 0.49 |
| 4 | 0.48 | 0.54 | **0.51** | 0.51 | 0.46 | 0.49 |

Table 5: Results by level for the two best models. L represents the level, *Prec* the precision, *Rec* the recall, and *F1* the F-measure. The best results are shown in bold.

Overall, our experimental results are consistent with the trends observed in previous studies, particularly the marginal advantage of monolingual models. Direct comparison, however, is not possible due to the significantly smaller test set used in the prior work (Rodríguez Rey and Garcia, 2025).

After discussing the results, they are then analyzed in conjunction with similar studies to provide a more comprehensive answer to the research questions.

- A1: For monolingual models and limited resources, increasing data by incorporating automatically translated data seems to be beneficial. However, for multilingual models, this strategy is only beneficial in some cases, and the margin of improvement is small. Furthermore, studies for similar tasks obtain slightly lower results with translated texts than with original texts (Rehan et al., 2023). Therefore, considering the time, effort, and computing costs involved, this strategy may not be worthwhile for multilingual models.

- A2: Fine-tuning models to classify texts into multiple classes remains worthwhile because the performance of generative models for this task is far from that of fine-tuned models, as shown by the results of this study and other similar works (Imperial et al., 2025). The same conclusion has been reached in other text classification tasks, such as classifying texts by genre or textual theme (Pungeršek et al., 2025). Furthermore, using language models with fine-tuning techniques achieves state-of-the-art results in numerous natural language understanding tasks (Devlin et al., 2019; Ngo and Parmentier, 2023). Many studies use this technique to classify texts according to their complexity since it is currently the most widespread technique in the field (Trokhymovych et al., 2024; Ribeiro et al., 2024; Coban et al., 2024; Vásquez-Rodríguez et al., 2022; Hernandez et al., 2022).

  However, other studies draw slightly different conclusions. They argue that the results obtained by fine-tuned models are competitive with those of generative models and with much shorter computation times (Kostina et al., 2025). It should also be noted that this study employed a standard methodology and that other approaches exist—such as using the

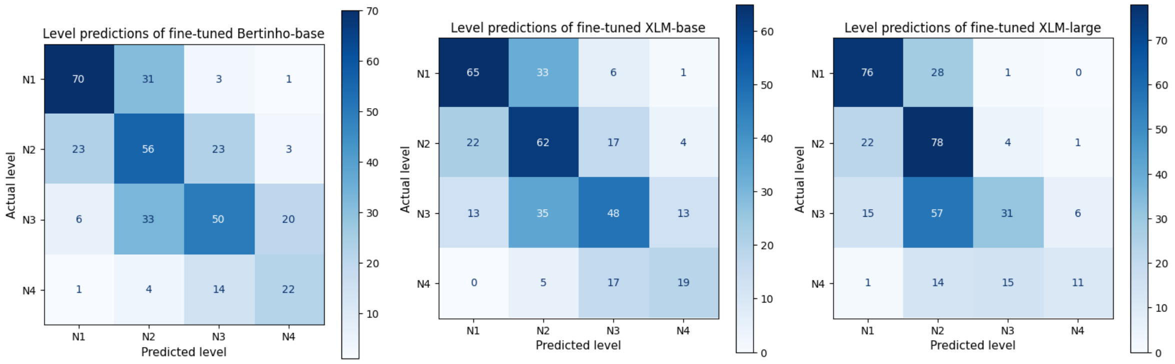


Figure 1: Distribution of predictions from the models with the best results on the test set.

same prompt translated into English, or even more advanced methods, such as using model-generated rubrics (Liu et al., 2024) or fine-tuning LLMs with LoRA (Yousefiramandi and Cooney, 2025)—which could yield better results.

Regarding the use of instructed or uninstructed models, the results obtained do not clearly support Hypothesis 2 (see Section 3) that instructed models perform better on this task. As for zero-shot and few-shot techniques, the results show that the few-shot technique, which includes several examples of the task, achieves better results.

## 7. Conclusions and further work

This paper have presented a comparative study of strategies for developing automatic text classifiers by readability level in Galician, as well as data augmentation approaches. The primary methods for developing classifiers involve using fine-tuned models (monolingual and multilingual) and large generative language models that are provided with instructions and examples of the task. Data augmentation techniques include using suitable task-specific datasets in other languages and automatically translating them into Galician.

To conduct the comparative study, datasets and models were selected and a series of experiments were carried out to answer the research questions. A baseline model combining linguistic features, vectors, and surprisal was established to evaluate the performance of the different models. Regardless of the type of model or data used, these models obtained similar results, with the best model achieving an accuracy of 0.51 using the translated dataset and a monolingual model.

Fine-tuned models consistently produce equal or superior results in both monolingual and multilingual experiments. Considering all the analyzed metrics, the best result is obtained with the Bertinho base model and translated data. Other models achieve slightly lower results, including both monolingual and multilingual models with both original and translated data. In contrast, generative models using a standard methodology, have shown lower performance compared to baseline models. In general, generative models provided with examples (few-shot) perform better than those provided only with instructions.

This study evaluated the results of different experiments and other similar works and concluded that, in the case of Galician, the strategy of data augmentation through machine translation of corpora designed for the task in other languages is beneficial when using monolingual models. For multilingual models, however, the margin for improvement may be minimal or nonexistent. Regarding generative models' performance in classifying Galician texts by readability level, it is concluded that, with a standard methodology, for now, they are not competitive with fine-tuned models.

Future work includes expanding the Galician readability corpus by increasing the number of texts per readability level and extending the range of levels covered. Further research may explore more sophisticated prompting strategies with LLMs and fine-tuning approaches beyond encoder-based architectures. These efforts aim to assess how data augmentation, cross-lingual transfer, and model design impact the automatic readability assessment of texts in Galician.

## 8. Limitations

Regarding generative models, the experiments were limited to the standard methodology of using two nearly identical prompts for the zero-shot and few-shot strategies. A more diverse approach would have involved testing prompts in different languages, such as Galician, Spanish, and English. Additionally, more advanced strategies could have been explored, such as model-generated rubrics, to better understand AI-specific reasoning compared to the human-defined criteria used in this study. Other strategies that could have been explored include fine-tuning via LoRA. Regarding the data, the evaluation corpus was relatively small (480 texts), and the translated corpus lacked manual supervision. Lastly, while we tested several architectures, using a wider variety of Transformers and LLMs could have strengthened the generalizability of the findings.

## 9. Acknowledgements

This work was supported by the European Commission (Project: iRead4Skills, Grant number: 1010094837, Topic: HORIZON-CL2-2022-TRANSFORMATIONS-01-07, DOI: 10.3030/101094837), by MCIU/AEI/10.13039/501100011033 (grants with references PID2021-128811OA-I00, CNS2024-154902, PID2024-161928OB-I00, and AIA2025-163322-C62), and by the Galician Government (ERDF 2024-2027: Call ED431G 2023/04, and ED431B 2025/16).

## 10. Lay Summary

Automatic tools that measure how difficult a text is to read are essential for language learning and making information accessible to everyone. While these tools are common for major languages, they are still rare for languages with fewer digital resources, like Galician. This study addresses this gap by testing different artificial intelligence methods to see which works best for classifying Galician texts by their complexity. Since the available native Galician dataset for this specific task is very limited in size, we used machine translation to create a larger set of training materials for the AI. We then compared two main strategies: training specialized models specifically for this task and using large, general-purpose AI models by giving them clear instructions and examples. Our results show that the specialized models are currently much more effective than the general AI models at accurately judging text difficulty in Galician. Furthermore, we found that using translated data significantly helps models that focus solely on Galician, whereas it has a limited impact on models that are already trained in multiple languages.

## 11. Bibliographical References

Onder Coban, Mete Yağanoğlu, and Ferhat Bozkurt. 2024. Domain effect investigation for bert models fine-tuned on different text categorization tasks. *Arabian Journal for Science and Engineering*, 49(3):3685–3702.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Unsupervised cross-lingual representation learning at scale. *CoRR*, abs/1911.02116.

Scott Andrew Crossley, David F. Dufty, Philip M. McCarthy, and Danielle S. McNamara. 2007. Toward a new readability: A mixed model approach. In *Proceedings of the annual meeting of the cognitive science society*, volume 29.

Edgar Dale and Jeanne Sternlicht Chall. 1948. *A formula for predicting readablility*. Bureau of Educational Research, Ohio State University.

Iria de Dios-Flores, Silvia Paniagua Suárez, Cristina Carbajal Pérez, Daniel Bardanca Outeiriño, Marcos Garcia, and Pablo Gamallo. 2024. CorpusNÓS: A massive Galician corpus for training large language models. In *Proceedings of the 16th International Conference on Computational Processing of Portuguese - Vol. 1*, pages 593–599, Santiago de Compostela, Galicia/Spain. Association for Computational Lingustics.

Felice Dell'Orletta, Simonetta Montemagni, and Giulia Venturi. 2014. Assessing document and sentence readability in less resourced languages and across textual genres. *ITL - International Journal of Applied Linguistics*, 165:163–193.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Rudolf Flesch. 1948. A readability formula in practice. *Elementary English*, 25(6):344–351.

Marcos Garcia. 2021. Exploring the representation of word meanings in context: A case study on homonymy and synonymy. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3625–3640, Online. Association for Computational Linguistics.

Gemma Team. 2025. Gemma 3.

José Ángel González, Ian Borrego Obrador, Álvaro Romo Herrero, Areg Mikael Sarvazyan, Mara Chinea-Ríos, Angelo Basile, and Marc Franco-Salvador. 2026. Iberbench: Llm evaluation on iberian languages. *Computer Speech & Language*, 96:101899.

Itziar Gonzalez-Dios, María Jesús Aranzabe, and Arantza Díaz de Ilarraza. 2018. The corpus of basque simplified texts (cbst). *Language Resources and Evaluation*, 52(1):217–247.

Naman Goyal, Cynthia Gao, Vishrav Chaudhary, Peng-Jen Chen, Guillaume Wenzek, Da Ju, Sanjana Krishnan, Marc'Aurelio Ranzato, Francisco Guzmán, and Angela Fan. 2022. The Flores-101 Evaluation Benchmark for Low-Resource and Multilingual Machine Translation. *Transactions of the Association for Computational Linguistics*, 10:522–538.

Aaron Grattafiori, et al., 2024. The Llama 3 Herd of Models.

G Harry and Mc Laughlin. 1969. Smog grading - a new readability formula. *The Journal of Reading*.

Nicolas Hernandez, Nabil Oulbaz, and Tristan Faine. 2022. Open corpora and toolkit for assessing text readability in French. In *Proceedings of the 2nd Workshop on Tools and Resources to Empower People with REAding DIfficulties (READI) within the 13th Language Resources and Evaluation Conference*, pages 54–61, Marseille, France. European Language Resources Association.

Carina Kauf and Anna A. . 2023. A better way to do masked language model scoring. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 925–935, Toronto, Canada. Association for Computational Linguistics.

Arina Kostina, Marios D. Dikaiakos, Dimosthenis Stefanidis, and George Pallis. 2025. Large language models for text classification: Case study and comprehensive review.

Yuxuan Liu, Tianchi Yang, Shaohan Huang, Zihan Zhang, Haizhen Huang, Furu Wei, Weiwei Deng, Feng Sun, and Qi Zhang. 2024. Calibrating LLM-Based Evaluator. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2638–2656.

Ion Madrazo Azpiazu and Maria Soledad Pera. 2019. Multiattentive recurrent neural network architecture for multilingual readability assessment. *Transactions of the Association for Computational Linguistics*, 7:421–436.

Joseph Marvin Imperial. 2021. BERT embeddings for automatic readability assessment. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pages 611–618, Held Online. INCOMA Ltd.

Joseph Marvin Imperial, Abdullah Barayan, Regina Stodden, Rodrigo Wilkens, Ricardo Muñoz Sánchez, Lingyun Gao, Melissa Torgbi, Dawn Knight, Gail Forey, Reka R. Jablonkai, Ekaterina Kochmar, Robert Joshua Reynolds, Eugénio Ribeiro, Horacio Saggion, Elena Volodina, Sowmya Vajjala, Thomas François, Fernando Alva-Manchego, and Harish Tayyar Madabushi. 2025. UniversalCEFR: Enabling open multilingual research on language proficiency assessment. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 9703–9755, Suzhou, China. Association for Computational Linguistics.

Kanishka Misra. 2022. minicons: Enabling flexible behavioral and representational analyses of transformer language models. *arXiv preprint arXiv:2203.13112*.

Ricardo Monteiro, Raquel Amaro, Susana Correia, Alice Pintard, Roser Gauchola, Michell Moutinho, and Xavier Blanco Escoda. 2024. iRead4Skills - Complexity Levels.

Duy Van Ngo and Yannick Parmentier. 2023. Towards sentence-level text readability assessment for French. In *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, pages 78–84, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Partha Pakray, Alexander Gelbukh, and Sivaji Bandyopadhyay. 2025. Natural language processing applications for low-resource languages. *Natural Language Processing*, 31(2):183–197.

Taja Kuzman Pungeršek, Peter Rupnik, Ivan Porupski, Vuk Dinić, and Nikola Ljubešić. 2025. State of the art in text classification for south slavic languages: Fine-tuning or prompting?

Andre Quispesaravia, Walter Perez, Marco Sobrevilla Cabezudo, and Fernando Alva-Manchego. 2016. Coh-Metrix-Esp: A complexity analysis tool for documents written in Spanish. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 4694–4698, Portorož, Slovenia. European Language Resources Association (ELRA).

Muhammad Rehan, Muhammad Shahid Iqbal Malik, and Mona Mamdouh Jamjoom. 2023. Fine-tuning transformer models using transfer learning for multilingual threatening text identification. *IEEE Access*, 11:106503–106515.

Eugénio Ribeiro, Nuno Mamede, and Jorge Baptista. 2024. Automatic text readability assessment in European Portuguese. In *Proceedings of the 16th International Conference on Computational Processing of Portuguese - Vol. 1*, pages 97–107, Santiago de Compostela, Galicia/Spain. Association for Computational Linguistics.

Sandra Rodríguez Rey and Marcos Garcia. 2025. Clasificación automática de textos por niveis de lecturabilidade: recursos e modelos para o galego. *Linguamática*, 17(2):33–56.

Sarah Schwarm and Mari Ostendorf. 2005. Reading level assessment using support vector machines and statistical language models. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 523–530, Ann Arbor, Michigan. Association for Computational Linguistics.

Johannes Sibeko. 2024. Harnessing google translations to develop a readability corpus for sesotho: An exploratory study. *Journal of the Digital Humanities Association of Southern Africa (DHASA)*, 5(1).

Qwen Team. 2024. Qwen2.5: A party of foundation models.

Mykola Trokhymovych, Indira Sen, and Martin Gerlach. 2024. An open multilingual system for scoring readability of Wikipedia. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6296–6311, Bangkok, Thailand. Association for Computational Linguistics.

Sowmya Vajjala. 2022. Trends, Limitations and Open Challenges in Automatic Readability Assessment Research. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 5366–5377.

Laura Vásquez-Rodríguez, Pedro-Manuel Cuenca-Jiménez, Sergio Morales-Esquivel, and Fernando Alva-Manchego. 2022. A benchmark for neural readability assessment of texts in Spanish. In *Proceedings of the Workshop on Text Simplification, Accessibility, and Readability (TSAR-2022)*, pages 188–198, Abu Dhabi, United Arab Emirates (Virtual). Association for Computational Linguistics.

David Vilares, Marcos Garcia, and Carlos Gómez-Rodríguez. 2021. Bertinho: Galician bert representations. *Procesamiento del Lenguaje Natural*, 66(0):13–26.

Rodrigo Wilkens, David Alfter, Xiaoou Wang, Alice Pintard, Anaïs Tack, Kevin P. Yancey, and Thomas François. 2022. FABRA: French aggregator-based readability assessment toolkit. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1217–1233, Marseille, France. European Language Resources Association.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.

Weihao Xuan, Rui Yang, Heli Qi, Qingcheng Zeng, Yunze Xiao, Aosong Feng, Dairui Liu, Yun Xing, Junjue Wang, Fan Gao, Jinghui Lu, Yuang Jiang, Huitao Li, Xin Li, Kunyu Yu, Ruihai Dong, Shangding Gu, Yuekang Li, Xiaofei Xie, Felix Juefei-Xu, Foutse Khomh, Osamu Yoshie, Qingyu Chen, Douglas Teodoro, Nan Liu, Randy Goebel, Lei Ma, Edison Marrese-Taylor, Shijian Lu, Yusuke Iwasawa, Yutaka Matsuo, and Irene Li. 2025. MMLU-ProX: A Multilingual Benchmark for Advanced Large Language Model Evaluation. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, 1513–1532.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, Jin Xu, Jingren Zhou, Jinze Bai, Jinzheng He, Junyang Lin, Kai Dang, Keming Lu, Keqin Chen, Kexin Yang, Mei Li, Mingfeng Xue, Na Ni, Pei Zhang, Peng Wang, Ru Peng, Rui Men, Ruize Gao, Runji Lin, Shijie Wang, Shuai Bai, Sinan Tan, Tianhang Zhu, Tianhao Li, Tianyu Liu, Wenbin Ge, Xiaodong Deng, Xiaohuan Zhou, Xingzhang Ren, Xinyu Zhang, Xipin Wei, Xuancheng Ren, Yang Fan, Yang Yao, Yichang Zhang, Yu Wan, Yunfei Chu, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zhihao Fan.

2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

Amirhossein Yousefiramandi and Ciaran Cooney. 2025. Fine-Tuning Causal LLMs for Text Classification: Embedding-Based vs. Instruction-Based Approaches. *arXiv preprint arXiv:2512.12677*.

A Ziyaden, A Yelenov, F Hajiyev, S Rustamov, and A Pak. 2024. Text data augmentation and pre-trained language model for enhancing text classification of low-resource languages. *PeerJ Computer Science*, 10:e1974.

## 12. Language Resource References

Alice Pintard, Thomas François, Justine Nagant de Deuxchaisnes, Sílvia Barbosa, Maria Leonor Reis, Michell Moutinho, Ricardo Monteiro, Raquel Amaro, Susana Correia, Sandra Rodríguez Rey, Marcos Garcia González, Keran Mu, and Xavier Blanco Escoda. 2024. iRead4Skills Dataset 1: corpora by complexity level for FR, PT and SP.

Sandra Rodríguez Rey and Marcos Garcia. 2025. Corlega: Corpus de Lecturabilidade en Galego.

## 13. Appendix A. Prompt used in LLM experiments

[Rol] Actúa como un lingüista experto en gallego, especializado en clasificación de textos por complejidad. Eres experto en clasificar textos en los niveles de complejidad definidos a continuación. Respondes a los textos de entrada que aparecen en el siguiente mensaje indicando claramente a qué nivel pertenecen.

[Objetivo] Tu objetivo es clasificar textos de entrada según su complejidad en cuatro niveles: tres niveles definidos y un cuarto nivel no definido que incluye textos más complejos que los definidos. Las clasificaciones deben ser claras y solamente en uno de los niveles. Los niveles se indicarán solamente con un número del 1 al 4, correspondiendo el 1 al nivel 1, el 2 al nivel 2, el 3 al nivel 3 y el 4 al nivel 4. Contarás con una serie de ejemplos de cada nivel para que entiendas mejor la tarea. Estas son las características de los diferentes niveles:

- Nivel 1: Incluye textos cotidianos, como mensajes cortos de correo o relacionados con las actividades diarias, letreros, listados, manuales, propaganda, formularios, materiales relacionados con sus intereses, documentos auténticos breves (billetes, entradas, cartas de restaurante, facturas, etiquetas, planos, embalajes, horarios, mapas... ) de uso muy frecuente, instrucciones sencillas, relatos fáciles, letras de canciones y poemas sencillos. Incluye los adverbios más comunes de cantidad (moito, pouco), tiempo (hoxe, agora, cedo), lugar (enriba, lonxe, preto), modo (ben, mellor, amodo), exclusión (só, soamente), afirmación y negación. Características Léxico-conceptuales: Nombres comunes y propios. Sufijo- iño/a. Valor de los afijos más frecuentes. Superlativo en- ísimo/a. Repertorio de vocabulario limitado a situaciones cotidianas y de especial interés. Características verbales: Verbos copulativos. Verbos regulares y semirregulares de uso frecuente (durmir, servir, espir). Irregulares más frecuentes (ir, facer, estar). Reflexivos y pronominales más comunes (chamarse, esquecerse, queixarse). Tiempos de indicativo en presente, pasado y futuro. Futuro hipotético. Perífrases más habituales: ir + infinitivo, estar/andar + gerundio, volver + infinitivo, comezar a + infinitivo, ter que + infinitivo. Características de estructura sintáctica: Coordinadas simples con e, mais, ou, pero. Oraciones temporales (cando) y comparativas y superlativas básicas. Afirmaciones, negativas, interrogativas con partículas y exclamativas básicas con partículas. Subordinadas simples con conjunciones más frecuentes: así que, porque, cando... Características de cohesión: Preposiciones y locuciones preposicionales más frecuentes. Uso de pronombres sólo con referente claro. Elipse sólo de elementos conocidos y muy claros. Marcadores de orden del discurso, de espacio y de tiempo más frecuentes (hoxe, agora, cedo, lonxe, preto, enriba...)
- Nivel 2: Incluye mensajes de carácter personal (SMS, correos, cartas) o de carácter social con frases rutinarias, cuestionarios sencillos, notas y mensajes relacionados con sus actividades de trabajo, estudio y ocio y textos sociales breves tipificados (para felicitar, invitar, aceptar/rehusar, agradecer, solicitar un servicio, pedir disculpas). Géneros textuales del nivel 1 y también páginas web, recetas, periódicos y revistas (titulares, noticias y anuncios) , horóscopos, cuentos y novelas cortos, anuncios de trabajo... Instrucciones sobre seguridad, aparatos de uso frecuente, alojamiento, menú de restaurante, vida académica, sanidad, posología de medicamentos... Información sencilla de formularios y documentos administrativos (Correos, Administración Pública, bancos, entidades académicas... ). Registro formal e informal. Adverbios más comunes de cantidad (moito, pouco), tiempo (hoxe, agora, cedo), lugar (enriba, lonxe, preto), modo (ben, mellor, amodo), exclusión (só, soamente), inclusión (até, mesmo, incluso), afirmación, negación y duda. Características léxico-conceptuales: Nombres comunes y propios. Interjecciones más frecuentes. Vocabulario de situaciones cotidianas y actividades habituales (por especial interés) y un vocabulario receptivo más amplio que permita comprender textos (sentimientos, meteorología, fauna y flora, alimentación, vida académica, salud, Administración, tiempo libre, actividades profesionales más habituales). Sufijo- iño/a. Afijos comunes relativamente frecuentes. Superlativo absoluto y relativo. Comparativos analíticos y sintéticos más frecuentes, forma meirande. Usos de ti, vós / vostede, vostedes. Características verbales: Verbos copulativos. Verbos regulares y semirregulares (durmir, servir, espir) e irregulares más frecuentes (ser, estar, facer, saber, querer, ir, vir, haber, poder, ter, deber, etc.). Reflexivos y pronominales más comunes (chamarse, esquecerse, queixarse, lavarse, vestirse). Tiempos de indicativo en presente, pasado y futuro. Futuro hipotético. Infinitivo, gerundio y participio. Presente de subjuntivo para expresión de deseos, sentimientos y reacciones (Oxalá veñas!/Que teñas sorte!/Sinto que marches). Imperativo para consejos, instrucciones y órdenes. Perífrases más habituales: estar, andar + gerundio, volver, comezar a, ter que, estar a/para, estar a/andar a + infinitivo, ter + participio, haber (de) + infinitivo. Características de estructura sintáctica: Coordinadas copulativas (e, mais, nin), adversativas (mais, pero, senón) y disyuntivas (ou...ou, nin...nin). Oraciones impersonales

sencillas. Afirmaciones, negativas (incluyendo doble negación), interrogativas con partículas, exclamativas e imperativas. Subordinadas simples con conjunciones más frecuentes: así que, porque, cando, ademais... Subordinadas temporales (cando, ao + inf) , causales, finales, condicionales y comparativas (máis...ca, meirande...ca). Características de cohesión: Preposiciones, locuciones preposicionales, conectores y enlaces más frecuentes. Uso de prons sólo con referente claro. Elipse sólo de elementos conocidos y muy claros. Marcadores de orden del discurso, de espacio y de tiempo más frecuentes y también entón, logo, por outra banda...

- Nivel 3: Incluye textos breves en lengua estándar relacionados con la vida cotidiana que informen sobre acontecimientos o transmitan opiniones, deseos u órdenes. También textos sencillos sobre temas de su ámbito académico o profesional. Entre otros: cartas de restaurantes, formularios, anuncios, folletos, visitas turísticas, hoteles, alquiler, cartas y correos, relatos, anuncios o artículos de la prensa sobre temas de actualidad, prospectos de medicamentos, folletos de divulgación y publicitarios... Características léxico-conceptuales: Nombres comunes y propios. Vocabulario de situaciones cotidianas, actividades de ocio, sentimientos, del ámbito laboral y de temas de interés general, como salud, accidentes, tecnología, medio natural, economía, sociedad, geografía, servicios... Características verbales: Voz pasiva. Tiempos de indicativo en presente, pasado y futuro, futuro hipotético y presente de subjuntivo de los verbos regulares y de los irregulares más comunes. Imperativo (también negativo) para consejos, instrucciones y órdenes. Perífrasis verbales frecuentes: poder + infinitivo (obrigación) , deber (de) + infinitivo (obrigación) , haber + inf. (obrigación) , ter + part. (aspecto) , andar + gerundio, ir + gerundio, estar para + infinitivo, poñerse a + infinitivo ), volver + infinitivo, empezar a + infinitivo, acabar de + infinitivo, deixar de + infinitivo, levar + gerundio, seguir + gerundio, dar + participio... Características de estructura sintáctica: Coordinadas copulativas (e, mais, nin) , adversativas (mais, pero, senón) y disyuntivas (ou...ou, nin...nin) , distributivas (ben... ben, ora... ora) y explicativas (é dicir, ou sexa, isto é). Usos de si: oraciones impersonales, pasivas y verbos reflexivos. Afirmaciones, negativas (incluyendo doble negación) , interrogativas con partículas (también indirectas) , exclamativas e imperativas. Subordinadas con subjuntivo e infinitivo. Subordinadas adjetivas y sustantivas, adverbiales (causales, finales, consecutivas, condicionales, temporales). Estilo indirecto. Características de cohesión: Preposiciones, locuciones preposicionales, conectores y enlaces más frecuentes. Uso de pronombres sólo con referente claro. Elipse sólo de elementos conocidos y muy claros. Marcadores de orden del discurso, de espacio y de tiempo más frecuentes y también entón, logo, por outra banda...

[Instrucción] Tienes que aprender los niveles definidos anteriormente y analizar los ejemplos que aparecen a continuación. Después, se te enviarán una serie de textos para que los clasifiques de uno en uno, por orden de aparición. Una aclaración: los textos contienen los caracteres “[SEP]”, que debes interpretar como saltos de línea en el texto. Cada texto aparece en una línea o fila. Estas son las tareas que debes realizar:

- Leer en profundidad el texto
- Utilizar las características de los niveles de complejidad descritos para analizar el texto, buscando, contando, señalando y extrayendo todas las que aparezcan en el texto
- Utilizar todas las características extraídas para determinar de qué nivel es el texto
- Responder con el nivel del texto basándote en los pasos anteriores, indicando solamente un número del 1 al 4

Estos pasos los tienes que repetir para cada texto.

[Contexto]

La información de los niveles con sus características para analizar aparecen en este mensaje, en la sección [Objetivo].

[Inputs esperados dentro de este mensaje]

- Niveles de complejidad y sus características
- Lista de textos para clasificar

[Formato y criterios de salida]

- Tabla csv con el nivel de cada texto de entrada, por orden.

[Restricciones]

- No inventes características que no estén en la descripción de características.
- No definas el nivel de un texto sin estar seguro de que es acorde.
- No respondas información que no sea el nivel del texto.

[Errores comunes a evitar]

- Inventar características
- No hacer el recuento de las características correctamente
- Inventar niveles

### 13.1. English translation of the prompt used in LLM experiments

[Role] You act as an expert linguist in Galician, specializing in classifying texts by complexity. You are an expert at classifying texts into the complexity levels defined below. You respond to the input texts that appear in the following message by clearly indicating which level they belong to.

[Objective] Your objective is to classify input texts according to their complexity into four levels: three defined levels and a fourth undefined level that includes texts more complex than those defined. Classifications must be clear and assigned to only one of the levels. The levels will be indicated only by a number from 1 to 4, where 1 corresponds to level 1, 2 to level 2, 3 to level 3, and 4 to level 4. You will be provided with a series of examples for each level to help you better understand the task. These are the characteristics of the different levels:

- Level 1: Includes everyday texts, such as short emails or messages related to daily activities, signs, lists, manuals, advertisements, forms, materials related to their interests, and brief authentic documents (tickets, admission tickets, restaurant menus, invoices, labels, maps, packaging, and schedules) in very frequent use, simple instructions, easy stories, song lyrics, and simple poems. Includes the most common adverbs of quantity (much, little), time (today, now, soon), place (above, far, near), manner (well, better, slowly), exclusion (only, solely), and affirmation and negation. Lexical-conceptual features: Common and proper nouns. Suffix -iño/a. Meaning of the most frequent affixes. Superlative -ísimo/a. Vocabulary limited to everyday situations and topics of special interest. Verbal features: Copulative verbs. Frequently used regular and semi-regular verbs (to sleep, to serve, to undress). Most common irregular verbs (to go, to do, to be). Most common reflexive and pronominal verbs (to be called, to forget, to complain). Indicative tenses in the present, past, and future. Hypothetical future. Most common periphrases: ir (to go) + infinitive, estar/andar (to be/to go) + gerund, voltar (to return) + infinitive, comezar a (to begin to) + infinitive, ter que (to have to) + infinitive. Syntactic structure features: Simple coordinating conjunctions such as “and,” “more,” “or,” and “but.” Temporal clauses (when) and basic comparative and superlative forms. Affirmative, negative, and interrogative sentences with particles, and basic exclamatory sentences with particles. Simple subordinate clauses with the most common conjunctions: so, because, when... Cohesion features: The most common prepositions and prepositional phrases. Use of pronouns only with a clear referent. Ellipsis only of known and very clear elements. The most common discourse, spatial, and temporal markers (today, now, soon, far, near, above...)
- Level 2: Includes personal messages (text messages, emails, letters) or social messages containing routine phrases, simple questionnaires, notes, and messages related to work, study, and leisure activities, as well as short, standard social texts (for congratulating, inviting, accepting/declining, thanking, requesting a service, and apologizing). Text types from Level 1, as well as web pages, recipes, newspapers, and magazines (headlines, news, and advertisements), horoscopes, short stories and novellas, job postings... Safety instructions, frequently used devices, housing, restaurant menus, academic life, healthcare, medication dosages... Simple information from forms and administrative documents (Post Office, Government, banks, academic institutions...). Formal and informal register. Most common adverbs of quantity (a lot, a little), time (today, now, soon), place (above, far, near), manner (well, better, slowly), exclusion (only, just), inclusion (up to, even, including), affirmation, negation, and doubt. Lexical-conceptual features: Common and proper nouns. Most frequent

interjections. Vocabulary for everyday situations and common activities (of special interest) and a broader receptive vocabulary that allows for understanding texts (feelings, weather, flora and fauna, food, academic life, health, government, leisure, most common professional activities). Suffix -iño/a. Relatively frequent common affixes. Absolute and relative superlatives. Most frequent analytical and synthetic comparatives, "meirande" (greater) form. Uses of "ti, vós / vostede, vostedes (you). Verb characteristics: Copulative verbs. Regular and semi-regular verbs (to sleep, to serve, to undress) and the most common irregular verbs (to be, to exist, to do, to know, to want, to go, to come, to have, to be able to, to have, to must, etc.). Most common reflexive and pronominal verbs (to be called, to forget, to complain, to wash, to get dressed). Indicative tenses in the present, past, and future. Hypothetical future. Infinitive, gerund, and participle. Present subjunctive for expressing wishes, feelings, and reactions (I hope you come!/Good luck!/I'm sorry you're leaving). Imperative for advice, instructions, and commands. Most common periphrases: estar/andar (to be/to go) + gerund, volver a (to return) + infinitive, empezar a (to begin to) + infinitive, ter que (to have to) + infinitive, estar a punto de (to be about to) + infinitive, ir a (to be going to) + infinitive, haber + participle, haber de (to have to) + infinitive. Syntactic structure features: Coordinating conjunctions (and, more, nor), adversative conjunctions (but, however, instead), and disjunctive conjunctions (either...or, neither...nor). Simple impersonal sentences. Affirmative, negative (including double negation), interrogative with particles, exclamatory, and imperative sentences. Simple subordinate clauses with the most common conjunctions: so, because, when, also... Temporal subordinate clauses (when, as + infinitive), causal, final, conditional, and comparative (more...than, "meirande" (greater)... than). Cohesion features: Prepositions, prepositional phrases, the most common connectors and linking words. Use of pronouns only with a clear referent. Ellipsis only of known and very clear elements. The most common discourse, spatial, and temporal markers, as well as "then," "later," "on the other hand," etc.

- Level 3: Includes short texts in standard language related to everyday life that report on events or convey opinions, requests, or instructions. Also includes simple texts on topics within their academic or professional field. Among others: restaurant menus, forms, advertisements, brochures, tourist information, hotels, rentals, letters and mail, stories, news reports or press articles on current events, medication leaflets, informational and advertising brochures... Lexical-conceptual features: Common and proper nouns. Vocabulary related to everyday situations, leisure activities, feelings, the workplace, and topics of general interest, such as health, accidents, technology, the natural environment, the economy, society, geography, and services... Verbal features: Passive voice. Present, past, and future indicative tenses, future subjunctive, and present subjunctive of regular verbs and the most common irregular verbs. Imperative (including negative) for advice, instructions, and commands. Common verbal periphrases: poder (can/may) + infinitive, deber (de) (must/should) + infinitive, haber de (to have to) + infinitive, ter (to have to) + participle, andar (to go) + gerund, ir (to go) + gerund, estar para (to be about to) + infinitive, poñerse a (to start/begin) + infinitive, volver a (to do again) + infinitive, empezar a (to begin to) + infinitive, acabar de (to have just) + infinitive, dejar de (to stop) + infinitive, levar (to have been) + gerund, seguir (to continue) + gerund, dar (to give) + participle. Syntactic features: Coordinating conjunctions (and, more, nor), adversative conjunctions (more, but, rather), disjunctive conjunctions (either...or, neither...nor), distributive conjunctions (either...or, or...or), and explanatory conjunctions (that is, in other words, that is to say). Uses of "si": impersonal, passive, and reflexive verb clauses. Affirmative, negative (including double negation), interrogative with particles (including indirect), exclamatory, and imperative. Subordinate clauses with subjunctive and infinitive. Adjectival and noun clauses, adverbial clauses (causal, final, consecutive, conditional, temporal). Indirect speech. Cohesion features: Prepositions, prepositional phrases, the most frequent connectors and linking words. Use of pronouns only with a clear referent. Ellipsis only of known and very clear elements. The most frequent discourse, spatial, and temporal markers, as well as then, later, on the other hand...

[Instructions] You must familiarize yourself with the levels defined above and analyze the examples provided below. You will then be sent a series of texts to classify one by one, in the order they appear. Note: The texts contain the characters "[SEP]", which you should interpret as line breaks in the text. Each text appears on a single line. These are the tasks you must perform:

- Read the text thoroughly
- Use the characteristics of the complexity levels described to analyze the text, searching for, counting, identifying, and extracting all those that appear in the text
- Use all the extracted characteristics to determine the text's level

- Answer with the text's level based on the previous steps, indicating only a number from 1 to 4

You must repeat these steps for each text.

[Context]

The information on the levels and their characteristics to be analyzed appears in this message, in the [Objective] section.

[Expected inputs within this message]

- Complexity levels and their characteristics
- List of texts to classify

[Output format and criteria]

- CSV table with the level of each input text, in order.

[Restrictions]

- Do not invent features that are not listed in the feature description.
- Do not assign a level to a text unless you are certain it is appropriate.
- Do not provide information other than the text's level.

[Common Mistakes to Avoid]

- Inventing characteristics
- Not counting characteristics correctly
- Inventing levels

## 13.2. Examples added at the end of the prompt for few-shot experiments:

[Ejemplos]

[Texto][Nivel]

['-Cal é o teu horario de traballo? [SEP] -Depende. Algúns días traballo moitas horas e outros case non traballo. O que si, sempre me ergo tarde, ás dez ou ás once. [SEP] Despois, almorzo e logo dou un paseo. Ao mediodía volvo á casa, xanto algo e vexo a tele un pouco. [SEP] Logo, baixo á rúa e collo o coche. [SEP] Ás veces traballo ata as nove ou as dez da noite. Cando acabo, vou á casa, preparo a cea e leo un pouco. Déitome á unha ou ás dúas máis ou menos. [SEP] Xaquín Pereira, 38 anos, taxista.'][1]

['Cabalgata e Recepción dos Reis Magos [SEP] venres, 5 xaneiro, 2024 - 17:00 [SEP] Praza do Obradoiro, [SEP] Santiago de Compostela [SEP] Cabalgata e Recepción dos Reis Magos [SEP] Os Reis Magos de Oriente percorrerán as rúas da cidade ata a praza do Obradoiro. [SEP] Ao rematar terá lugar a tradicional recepción no Pazo de Raxoi para todos os nenos e nenas que queiran achegarse ás súas Maxestades. [SEP] PERCORRIDO PREVISTO: [SEP] Saída ás 17:00 horas desde a Praza da Mercé e desfile pola avenida de Ferrol, Frei Rosendo Salvado, praza Roxa, República de El Salvador, rúa do Hórreo, praza de Galicia, rúa da Senra, Porta Faxeira, rúa do Vilar, praza das Praterías, travesa de Fonseca e praza do Obradoiro. A recepción está prevista para as 19.00 horas.'][1]

['Con esta receita saen uns 12 Sapos da Limia xeitosos. Se queredes que vos saian máis, só tedes que multiplicar os ingredientes. E non vos preocupedes polo sabor a cervexa, que lles queda de vicio! [SEP] Ingredientes [SEP] 85 ml de cervexa [SEP] 125 ml de leite [SEP] 1 ovo [SEP] 3 culleradas de azucre [SEP] 1 belisco de sal [SEP] 230 gramos de fariña triga [SEP] 20 gramos de manteiga derretida [SEP] Aceite de xirasol para fritir [SEP] Azucre para rebozar sapos [SEP] Preparación [SEP] Imos mesturando os ingredientes por orde de aparición na listaxe. [SEP] Deixamos que repouse a masa media hora fóra da neveira. [SEP] Quentamos o aceite de xirasol nun caciño e imos botando culleradas de masa. Aquí a dica é botar toda a masa de golpe. Eu fíxeno coa culler dos xeados. [SEP] Rebozamos os sapos en azucre e xa estarían!'][1]

['As festas de Ribeira [SEP] Este ano as festas en honra a santa Uxía contarán coa presenza das dúas orquestras de referencia no ámbito musical galego. A orquestra Panorama e a París de Noia animarán o día central das festas de Ribeira, desde as primeiras horas da noite ata ben entrada a madrugada. Non só está garantida a calidade musical, senón tamén o espectáculo, a diversión e as ganas de bailar ata moi, moi tarde. Ademais, a comisión de festas agasallará a todos os asistentes coa tradicional queimada ao finalizar a verbena. Para que a diversión non decaia, agosto comeza en Ribeira!'][1]

['SANTIAGO DE COMPOSTELA [SEP] TAPAS [SEP] Para tomar tapas, o máis recomendable é facelo na zona vella, nas rúas do [SEP] Franco e da Raíña, onde están todos os bares con especialidades galegas: polbo á feira, empanada, pementos de [SEP] Padrón, mariscos... Algúns dos máis famosos son O Orella, onde podes tomar orella de porco, o SantYago, [SEP] Os Caracois, o Central, todos eles na rúa da Raíña. Son bares moi animados onde se mestura moita xente nova, estudantes, peregrinos e turistas. [SEP] ONDE COMER [SEP] A Barrola, con comida galega de boa calidade e a bo prezo; o Enxebre, ao lado da catedral; El Pasaje, con especialidade en mariscos e carnes; [SEP] A Alameda, con especialidade en empanada; a Casa Marcelo, comida galega moderna, ten unha estrela [SEP] Michelin.'][1]

['Do Alto do Acevo a Castroverde [SEP] Distancia: 45,1 km [SEP] Nos seus primeiros metros, o Camiño Primitivo atravesa unha fraga con carballos, bidueiros e castiñeiros. Esta primeira etapa cruza unha contorna de montaña, na que se coroan diferentes altos, como o da Fontaneira ou o da Vacariza. Nos vales, os prados e pastizais engaden textura á paisaxe. E entre a vexetación cómpre sinalar os piñeirais, tanto da especie do país como as repoboacións de piñeiro silvestre. [SEP] 1. Alto do Acevo [SEP] 43°8'49.3"N 6°58'24"W [SEP] O Acevo, a máis de mil metros de altitude, recibe o nome polos acivros que hai nesta zona. O gando doméstico que pastorea nos montes emprega as acivreiras para refuxiarse do frío no inverno e da calor no verán. Tamén acubillan aves como o paporroibo, o gaio ou o merlo común. [SEP] 2. Conxunto de acivros [SEP] 43°8'36.8"N 6°58'9.9"W [SEP] O acivro é un arbusto con follas perennes de cor verde escura. É resistente ao frío e medra devagar. Florece durante a primavera e os seus froitos, mantenza da fauna silvestre, maduran no seguinte inverno, presentando unha cor vermella e brillante. A súa madeira é dura e resistente. [SEP] 3. Alto da Fontaneira [SEP] 43°2'15.2"N 7°11'45.5"W [SEP] A súa altitude, a 936 metros, e a súa situación no interior de Galicia, fan que o clima deste lugar sexa de montaña. Caracterízase por ter unha temperatura media baixa, polo que son frecuentes as xeadas. Son abundantes tamén as precipitacións, en forma de neve no inverno. [SEP] 4. Fraga de Estornín [SEP] 43°1'44.11"N 7°12'32.96"W [SEP] As fragas son extensións de monte, xeralmente de difícil acceso, poboadas por árbores caducifolias. A de Estornín ten carballos, pradairos, capudres, teixos, bidueiros e castiñeiros. No seu interior habitan mamíferos salvaxes abundantes na zona, como lobos, corzos, xabarís ou raposos.'][2]

['Detida unha parella na Coruña que se fixo pasar por policía para roubar un móbil e unha carteira [SEP] A Policía Nacional detivo a un home e a unha muller que se fixeron pasar por policías para roubar un teléfono móbil e unha carteira na Coruña, ademais doutros roubos. Segundo informou o Instituto Armado, o primeiro dos roubos ocorreu na zona de Oza, no mes de agosto, cando abordaron a unha persoa que camiñaba soa pola rúa. [SEP] A muller, esgrimindo unha navalla, esixiulle que lle entregase as súas pertenzas. Así, conseguiron roubarlle o teléfono móbil e a súa carteira, que contiña diñeiro en efectivo e cartóns bancarios. Días máis tarde, ao redor das 7.00 horas da mañá, cometeron outro roubo na rúa Costa Rica cando a vítima ía coller un taxi e é sorprendida pola parella. [SEP] A parella fíxose pasar por policías e argumentaron que realizaban un control de drogas. Con esta escusa, empuxaron á vítima contra unha parede e subtraéronlle o seu teléfono móbil e a súa carteira para, posteriormente, abandonar o lugar. Nos dous casos, nas horas posteriores aos roubos, realizaron varios cargos cos cartóns bancarios subtraídos en diferentes establecementos da cidade.'][2]

['Falece o xornalista Pepe Seijo [SEP] Seijo, de 57 anos, estaba traballando nos informativos cando faleceu de maneira repentina. [SEP] Unha das voces máis características das ondas radiofónicas en Lugo, Pepe Seijo, faleceu esta segunda feira de forma repentina mentres traballaba no informativo diario de Radio Lugo-Cadena SER, segundo informa a TVG. O xornalista, de 57 anos de idade, entrou en directo no primeiro boletín da xornada, mais no segundo xa non interveu. [SEP] Seijo naceu en Bilbao e pasou a súa infancia en Euskadi. Seguidor apaixonado do Athletic de Bilbao, como se recoñecía el mesmo en calquera conversa, chegou a ser presidente da peña luguesa deste equipo. [SEP] Décadas de traballo ligado ás ondas servíronlle para converterse nunha voz de referencia na cidade e arredores, que perdeu no principio do ano outro dos seus xornalistas radiofónicos máis coñecidos, Arcadio Silvosa.'][2]

['Vigo prenderá as luces de Nadal o 24 de novembro ao que suma unha árbore de 44 metros, "a máis importante do mundo" [SEP] O alcalde de Vigo, Abel Caballero sinalou ese día como o momento en que acenderá "o Nadal de todo o planeta". [SEP] O alcalde de Vigo, Abel Caballero, confirmou este xoves que o acto de aceso da iluminación do Nadal terá lugar finalmente o 24 de novembro ás 20,30 horas, momento en que se acenderá "o Nadal de todo o planeta". [SEP] Así o trasladou en declaracións aos medios nunha visita á Porta do Sol, onde este xoves comezaron os traballos de instalación do gran abeto luminoso que, este ano, alcanzará os 44 metros de altura. "A árbore máis importante do mundo é o de Vigo", presumiu o rexedor, que explicou que a estrutura contará cunha gran estrela na súa cúspide, cunha lonxitude total de 19 metros.’][2]

[’Os cinco preceptos nas ensinanzas de Buda [SEP] Como toda relixión, o budismo conta con preceptos básicos que deben seguirse con rectitude. En total, só hai cinco, pero abarcan áreas importantes da vida. Os preceptos de Buda son "Non mates", "Non roubes", "Non abuses do sexo" e "Non consumas drogas nin alcohol". Entende a continuación a razón de cada un. [SEP] Non mates [SEP] É posible que toda relixión, filosofía ou doutrina teña en conta esta lei. As ensinanzas de Buda van un pouco máis aló que outras tradicións, porque cando di non mates -porque formas parte do todo e cometendo tal acto estás a facerte dano- tamén está a falar de animais, como a galiña, o boi ou ata unha formiga. [SEP] Non roubes [SEP] Se non queres o que pertence aos demais e estás satisfeito cos teus logros, xa vas por bo camiño. Pero aínda así, o budismo subliña a idea de que non se debe roubar, aínda que sexa o lugar de alguén na fila, froito do esforzo intelectual ou físico de alguén, ou mesmo de obxectos. [SEP] Non abuses do sexo [SEP] O sexo é absolutamente natural e moi ben visto no budismo, porén non deixa de ser un intercambio de enerxía e calquera exceso é visto de forma atenta polas ensinanzas de Buda. Polo tanto, é importante manter o acto sexual saudable e como complemento da túa vida, non como foco das relacións. [SEP] Non Consumas Drogas nin Alcol [SEP] Mantén a túa mente activa e sempre en plenitude, observar o momento presente é fundamental para lograr chegar a Magga, é dicir, a fin do sufrimento. Por outra banda, o uso de estupefacientes -legalizados ou non- altera o funcionamento do cerebro e, polo tanto, o seu uso non se recomenda no budismo.’][3]

[’O CLIENTE, A NOSA PRIORIDADE [SEP] Sempre ao seu servizo [SEP] O noso obxectivo é ofrecer aos clientes a mellor calidade aos mellores prezos, cun bo servizo e atención en centros de proximidade. Ademais, contamos cunha ampla variedade en frescos, produtos de marcas líderes e marca Froiz. O cliente está no centro de todas as decisións: "Sempre ao seu servizo", tentando dar resposta ás súas demandas e necesidades. [SEP] FROIZ [SEP] Nós [SEP] Somos unha empresa dedicada a venda por xunto e polo miúdo de produtos frescos, alimentación, adega e droguería, a través da nosa rede comercial formada por supermercados, hipermercados, cashcarry e tenda en liña. [SEP] A nosa actividade desenvólvese en España e Portugal. En España, estamos presente nas comunidades autónomas de Galicia, Castela e León, Castela-A Mancha e Comunidad de Madrid. A empresa sitúase entre as 20 primeiras empresas do sector a nivel nacional e forma parte da central de compras Euromadi. [SEP] COMPROMETIDOS COS NOSOS VALORES [SEP] O noso obxectivo empresarial é ofrecer aos clientes a mellor calidade aos mellores prezos, cun bo servizo e atención en centros de proximidade. [SEP] CALIDADE [SEP] OS MELLORES PREZOS [SEP] SERVIZO E ATENCIÓN [SEP] FORNECIDO AMPLO [SEP] PROXIMIDADE E COMODIDADE’][3]

[’Artigo 1. Obxecto e finalidade [SEP] 1. Esta orde ten por obxecto fixar as bases reguladoras do Programa para a promoción do emprego autónomo en Galicia (TR341D) e proceder á súa convocatoria para o ano 2024. [SEP] 2. A finalidade deste programa é a concesión de 2 liñas de axuda económica a aquelas persoas desempregadas que pretendan desenvolver a súa actividade empresarial ou profesional en Galicia como traballadoras autónomas ou por conta propia, para facer fronte aos distintos gastos xerados no comezo e mantemento da súa actividade laboral. [SEP] 3. Ao abeiro desta orde subvencionaranse as altas na Seguridade Social ou en mutualidade de colexio profesional e os gastos de mantemento da actividade, que, cumprindo os requisitos e condicións establecidas nela, se formalicen desde o 30 de setembro de 2023 ata o 29 de setembro de 2024, ambos inclusive. [SEP] Establécense 2 liñas de axudas: [SEP] a) Liña 1: unha axuda para o inicio da actividade económica e mantemento do emprego como persoa traballadora autónoma. [SEP] b) Liña 2: unha axuda para sufragar os gastos do mantemento da actividade por conta propia, equivalente a 12 meses da contía da cota reducida regulada no apartado 1 do artigo 38.ter da Lei 20/2007, do 11 de xullo, do Estatuto do traballo autónomo. [SEP] Nesta convocatoria incorpórase o establecemento de métodos de custos simplificados conforme ao disposto no Regulamento (UE) 2021/1060 do Parlamento Europeo e do Consello, do 24 de xuño de 2021.’][3]

[’OS HEAVIES E AS POETAS [SEP] Ao remate dun recital, un xornalista espétame: "Os heavies e os poetas sodes os últimos inocentes deste mundo". Nunca tal eu pensara, e iso que sempre me prestaron moito os heavies. Inspírame confianza a xente que responde á dureza das cousas armándose de dureza, desde o nome até o coiro. Teñen a sensibilidade dos que precisan responder ás agresións.

Teñen conciencia grupal. Non son submisos. Para os demais, cadeas e reloxos. Eles non van á moda, por iso teñen aínda o corazón de ouro. Os poetas, en cambio, habiamos levar máis ferro ao corazón. [SEP] Parada como quedei despois daquela arroutada, non fun quen de responder que non debería haber cousa menos inocente que un poeta. O profesor Antón Figueroa, referíndose ao campo cultural, fala adoito do "mito da inocencia". Os mitos dan lugar aos prexuízos. Os prexuízos impiden coñecer. Alguén di "leliadoura", e o lector, con bo tino, conclúe "plátano es". Os poetas adoitan deixar as frases en suspenso e os poemas, xa se sabe, son cousa moi ambigua. O plátano vén das Canarias. En Canarias é unha hora menos. En Galicia perdemos luz por culpa do cambio horario. (Isto non é poesía.) [SEP] Conclusión: a cultura come e dá de comer en boa medida grazas á implantación social do mito da inocencia. Que boa xente, os artistas. E os poetas, eses xa son o máximo. [SEP] Escriben con metáforas. Eles, menudo estilo. E elas, tan estilosas. [SEP] Hai moitos xeitos de perder a inocencia. Todos son bos. Algúns mecanismos de lexitimación son tan visíbeis que case non é preciso denuncialos. En cambio, adóitase ignorar a presenza efectiva (e por sutil, se cadra, dobremente efectiva) de modos de intervención máis inmediatos. Como se escribir publicamente non fose xa un xeito de exercer o poder. [SEP] A linguaxe tamén é violencia. Sempre que se abre un turno de palabra, outra persoa deixa de falar. A alternativa, quizais: non escribir para que nos escoiten, senón para que nos respondan. Non creo que o poder sexa malo de seu, e Foucault demostrou que vén de abaixo, hipótese que convida a exercelo até arriba. Non somos inocentes, as poetas, e non é mala cousa. A vida é complicada, e a vida cultural non é unha excepción. [SEP] Todos temos intereses, pero non todos os intereses son iguais. Quizais cumprise comezar por distinguir entre os públicos e os privados. Entre os persoais e os colectivos. [SEP] Para poder quedar, hai que asumir que estamos. Que non se nos xulgue pola presenza ou ausencia de estratexias, senón polo sentido das nosas estratexias. O que facemos ou deixamos de facer non será cabal na medida en que careza de propósito, senón na medida en que o seu propósito sexa máis ou menos cabal. [SEP] Deámonos o luxo de ser obxectados publicamente polo que procuramos, e non só polo que facemos e dicimos. As cousas moitas veces non son o que parecen. Para poder saír do reino da opinión, habería que comezar por dicir, canda Montaigne, "o que eu opino non dá a medida do mundo, senón a medida do meu entendemento". O mundo é ancho e alleo. Do entendemento, en poucos casos podemos dicir que é propio. Fagamos o posíbel por que non sexa estreito.'][4]