Leveraging LLMs for Cross-Locale Adaptation: a Workflow Proposal on Spanish Variants

Vera Senderowicz Guerra
Welocalize
vera.senderowicz@welocalize.com

Abstract

Localization strategies often vary significantly across languages, but the necessity of developing entirely separate approaches for closely related language variants remains debatable. This paper investigates the potential of streamlining the development process of localization strategies across Spanish locales. Leveraging Large Language Models, prompting techniques, and specialized linguistic resources, we explore methods for adapting a chosen baseline translation —produced by a Neural Machine Translation engine and post-edited by professional linguists— into region-specific variants. Focusing on transformations from Latin American Spanish into Mexican and Argentine Spanish, we examine vocabulary, terminology, grammar, and stylistic differences. Our findings suggest that building from a high-quality baseline and applying a modular, mostly automated adaptation process can efficiently address locale-specific divergences. While this approach reduces the need for manual intervention, human linguistic review remains essential, especially to refine stylistic nuances.

1 Introduction

Many international enterprises operating in diverse markets worldwide translate their content into multiple languages and localize it to the specific variants spoken by their target audiences. Despite the overarching goal of effective engagement, localization strategies can vary significantly between languages and even among different variants of the same language, due to factors such as translation volume, data availability, audience size and potential clients in each region, with the ultimate objective being to choose the most efficient and best suited solution for each market (Dunne and Dunne, 2011).

In this paper, we investigate the extent to which localization strategies can be streamlined for different variants of the same language. We propose a standardized workflow based on a common, human reviewed Neural Machine Translation (NMT) root, and a set of optional AI-powered post-editing steps that utilize Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), prompting techniques, and language resources. These steps are designed to address the divergences of each variant from the designated base and to make the necessary adjustments for adapting that base to different locales.

2 Experimental settings

This study is based on proprietary bilingual datasets provided by a commercial client in the entertainment industry. The source language is English (EN) —reflecting its centrality in both the client's operations and global markets— and the data comprise user interface segments and marketing copy. This content type is well-suited for the experiment because it constitutes a relatively low-risk domain, where minor regional inaccuracies are unlikely to cause significant consequences, and because it presents distinct localization challenges, as it often demands cultural specificity and audience engagement over linguistic neutrality. The same datasets are used throughout all stages of the experiment; they cannot be publicly released due to confidentiality agreements.

We focus on Spanish (ES), a language with numerous regional variants spoken in strategically important markets. Beyond its commercial relevance, ES offers a compelling case for this study due to the diversity of its variants across lexical, grammatical, and stylistic dimensions, each potentially requiring different adaptation strategies. The study examines three specific locales: Latin American (ES-LA), Argentine (ES-AR), and Mexican

Proceedings of Machine Translation Summit XX Volume 2, pages 58–66, June 23-27, 2025

Spanish (ES-MX). Due to the involvement of professional human translators in both comparison and evaluation tasks, the datasets are relatively small: 1637 segments for EN>ES-LA, 1635 for EN>ES-AR, and 1624 for EN>ES-MX. For cross-variant comparisons, only overlapping segments were retained: 1567 between ES-LA and ES-AR, 1495 between ES-AR and ES-MX, and 1474 between ES-LA and ES-MX. The ES-LA variant used is a commercially standardized form designed to be broadly accessible in pan-regional contexts when full regional localization is not feasible. The selection of only these three variants was guided by strategic relevance, data availability, and time constraints. While the approach is designed for potential reuse, its applicability to other locales and/or languages requires further specific testing and validation. It should also be noted that this approach is not intended for direct application to terminology-intensive domains such as legal or medical translation, which demand domain expertise, stricter quality controls, and accommodate more complex patterns of locale-specific terminology variation.

All LLM-based evaluations were conducted using OpenAI's GPT models, selected for their accessible fine-tuning capabilities (OpenAI). We chose GPT-4o mini due to its strong performance in automatic post-editing (Raunak et al., 2023) and its cost-efficiency relative to other OpenAI models (OpenAI, 2024). While we hypothesize that similar outcomes could be achieved with alternative models from other providers with minimal prompt adjustments (Uguet et al., 2024), model comparison was beyond the scope of this study, which focused on process development rather than tool benchmarking.

Prompting workflows were limited to three iterations per process and locale, following "Green AI" principles (Schwartz et al., 2020). Fine-tuning used no more than 55 examples per task. All experiments ran on 4 CPUs (Core i3-10350K) over 10h 30min, with an estimated carbon footprint of 276.37 gCO_2e (1.62 kWh), equivalent to 0.30 tree-months in Spain (calculated using Lannelongue et al., 2021).

3 The Spanish variations

To identify effective methods for transforming one ES locale into another, we first needed a clear understanding of what those transformations would entail. To this end, we conducted a contrastive linguistic analysis (Bennett, 2002; Ke, 2019) on reference translations from EN into ES-LA, ES-MX and ES-AR, all produced by the same NMT engine and post-edited by professional native linguists. This analysis led to the identification of the four categories of cross-locale divergences described below:

- **Terminological differences (client-specific terminology).** A subcategory of lexical changes, these terminological differences pertain to terms that primarily reflect the client's specific products and/or services and their preferred presentation to the target audience, rather than intrinsic characteristics of the ES variant itself.

- **Vocabulary differences (non-client-specific vocabulary).** Words and constructions that are preferred over others in different regions. Also a subcategory of lexical changes, these preferences are not dependent on the client's content but rather on the specific culture to which the content belongs. These preferences may include verbs (e.g., "regresar" is preferred over "volver" in ES-MX), nouns (e.g., "mamadera" is more commonly used than "biberón" in ES-AR), adjectives (e.g., while "small" tends to be translated as "chico" in ES-AR, "pequeño" is preferred in ES-LA), adverbs (e.g., "después" and "todavía" are more widely used in ES-AR than their corresponding "luego" and "aún"), and even different types of constructions (e.g., when used to convey a sense of duty, "tener que" is preferred over "deber" in ES-AR).

- **Grammatical differences.** While not all variants differ in this aspect, it is one of the most determining factors for recognizing ES locales: while certain words or terms might seem out of place if used in the context of a locale they don't belong to, verbs and pronouns conjugated according to the grammar rules of a different locale can lead to the entire text being identified as belonging to it. The primary difference usually lies in the second person: in this case, ES-LA and ES-MX don't present any differences, but ES-AR follows the "vos" conjugation ("vos amás", "vos querés", "vos partís"), instead of the widely used "tú" ("tú amas", "tú quieres", "tú partes").

- **Style differences.** The most complex category, it concerns how utterances sound “natural” within the cultural and communicative norms of each locale. Unlike grammar or terminology, style is less prescriptive: the rules governing it are highly context-dependent, often implicit, and nearly impossible to codify exhaustively. Additionally, style is shaped by overlapping factors such as client preferences, domain conventions, and regional usage, making style adaptation more intricate than other types of linguistic adjustment. Given this variability, it is challenging to provide universal examples, but some illustrative cases include the preference for the periphrastic future ("vas a venir") over the simple future ("vendrás") in ES-AR, and the use of constructions starting with "que lo" ("¡Que lo disfrutes!") instead of the imperative ("¡Disfrútalo!") in second person phrases expressing the speaker’s wishes.

While this categorization is based on linguistic criteria, its primary purpose is to group elements according to the similarity or compatibility of the rules governing their transformation, thus enabling a shared adaptation approach.

3.1 Deciding the baseline locale

It was necessary to determine which locale would serve as the baseline for transformations. In this context, *baseline* does not imply neutrality, but rather refers to the more “in-between” variant, the most practical starting point for adaptation. To define it, we compared the reference samples mentioned above using two of the most widely recognized —and most commonly requested by clients— machine translation quality metrics that assess the distance between a hypothesis and a reference translation: BLEU (Papineni et al., 2002) and Levenshtein Edit Distance (Levenshtein, 1966), normalized by the number of characters in the MT output, as shown in Table 1 below.

ES locales	BLEU	PE Distance
LA-AR	84.53	7%
LA-MX	92.69	5%
MX-AR	83.84	6%

Table 1: Distance between ES samples measured by BLEU and Levenshtein Edit Distance.

The first significant observation from Table 1 is that the metrics support the primary hypothesis of our experiment: if minimal editing effort is required to convert one locale to another, all locales are relatively “close”, which suggests that a strategy merge would not only be feasible but also sensible. Secondly, the results indicate that ES-AR might not be the best suited baseline candidate, as it is the most divergent from the other two locales. Additionally, it exhibits all four types of differences described when compared to both ES-MX and ES-LA, while these only display terminological and stylistic differences, which is reflected in their high similarity scores. Since the metrics indicate that both ES-MX and ES-LA are similarly suitable, we have chosen the latter as the baseline locale, based on our linguistic assessment: being a commercially constructed convention, it is better attained through human post-editing of NMT output following client-specific guidelines, as vocabulary and style-related uncertainties would be likely to arise during the adaptation process, with no underlying language community to inform such decisions beyond the client’s specifications.

4 Adaptations

After defining the baseline locale, we proceeded to develop an automatic post-editing method for each of the previously defined categories of differences. We adopted a segment-level approach, iterating through segment pairs to individually perform automatic post-editing on each of them.

4.1 Terminology

To adapt client-specific terminology, we used a glossary stored in a CSV file, with EN source terms in the first column and corresponding terms for each target locale in subsequent columns. The replacement logic was as follows: when an EN term from the glossary appears in the EN source segment and the ES-LA term is present in the target segment, it is replaced with the appropriate ES-MX or ES-AR term based on the locale. If the ES-LA entry is missing in the glossary, we verify that the ES-MX or ES-AR term corresponding to that EN entry is present in the ES target segment.

While Regular Expressions (RegEx) efficiently identify character patterns for checking compliance with the conditions described in the replacement logic above (Chapman and Stolee, 2016), their contextual limitations make replacement challenging due to the morphological richness of Spanish (Moreno-Sandoval and Goñi-Menoyo, 2002).

Many glossary entries require a context-aware insertion into the target segment, in a manner that aligns them with any word sharing the same referent. To address this, we used LLMs, which excel at context-dependent tasks (Qureshi et al., 2024).

We combined the generative capabilities of LLMs with RegEx's pattern recognition through a Term-Augmented Generation (TAG) technique inspired by the work of Sara Zanzottera for the 2024 AMTA Tutorial Day (Zanzottera, 2024). Instead of loading the entire glossary for each segment, TAG retrieves only relevant entries, which are inserted into a "Translation Guide" and prompted to the LLM along with general instructions for terminology replacement. The final instructions were refined iteratively based on output errors. Templates of the prompts used are provided in Appendices A, B, and C.

4.2 Vocabulary

Like terminology, vocabulary replacements often require morphological adaptation to remain grammatical, so we followed a similar approach to that described in Section 4.1. In the long term, it would be feasible to create and maintain an ES cross-locale vocabulary table for reuse in various projects within the same content type. The contrastive analysis revealed differences between ES-LA and ES-AR, but not between ES-LA and ES-MX. Due to the limited number of entries, we prompted the full list without TAG. As the table grows, the process could mirror that of Section 4.1, minus the need to retrieve the EN term. Additionally, some entries require instructional notes to guide replacements based on context. For example, "deber" changes to "tener que" in ES-AR, unless used in its reflexive form, which expresses causal relationships rather than obligations or instructions in all ES variants. A sample prompt used for ES-AR is provided in Appendix D.

4.3 Grammar

LLMs are exposed to large sets of multilingual data and have the potential to process context and therefore appropriately conjugate words according to locale-specific grammar rules (Penteado and Perez, 2023; Uchida, 2024).

Table 2 shows the distance increase between the reference ES-AR translation and the baseline translation after asking GPT-4o mini to adapt the latter's grammar to ES-AR rules using a zero-shot and a few-shot approach (original distance metrics are in Table 1). The zero-shot prompt is included in Appendix E. Most errors were due to limited recall and issues with correctly applying the appropriate conjugations: many verbs and pronouns were incorrectly pluralized or converted into the first person instead of being adapted to the "vos" conjugation. Furthermore, even when verbs were correctly adapted, surrounding pronouns and adjectives were not always adjusted accordingly.

Prompting approach	BLEU	PE Distance
Zero-shot	81.20	9%
Few-shot	83.24	8%

Table 2: Quality metrics of ES-LA into ES-AR grammatical adaptations performed by GPT-4o mini.

Building on our previous research (Senderowicz, 2024) —which demonstrated that fine-tuning is particularly effective for grammatical conjugation adaptations— we fine-tuned GPT-4o mini for ES-LA to ES-AR grammar transformation. To enhance the model's capabilities beyond what few-shot prompting could achieve, we followed OpenAI's fine-tuning procedures (OpenAI; OpenAI, 2023). We constructed training and validation sets featuring a range of transformation examples drawn from the generic model's most significant errors, targeting the most challenging structures. To promote precision and avoid over-editing, we also included examples requiring no change. We conducted three fine-tuning iterations, evaluating performance after each and incorporating new examples that mirrored grammatical patterns in previously mishandled cases.

4.4 Style

As stated above, style is the most nuanced aspect of language, shaped by tone, register, and cultural norms, and rarely governed by fixed rules that allow for a single "correct" choice. In fact, our linguistic review showed that many stylistic differences across ES variants required no editing; not because style is minor, but because multiple renderings were equally appropriate within the client's context. This underscores the need for human evaluation: style's highly subjective and context-sensitive nature makes it especially difficult to automate. For this reason, we chose not to automate style adaptation, considering the process successful if it addresses grammar and terminology while leaving stylistic choices to human reviewers. This decision preserves style as a domain for expert input and allows linguists to focus on high-impact, creative work tied to brand voice and communicative intent.

4.5 Final workflow

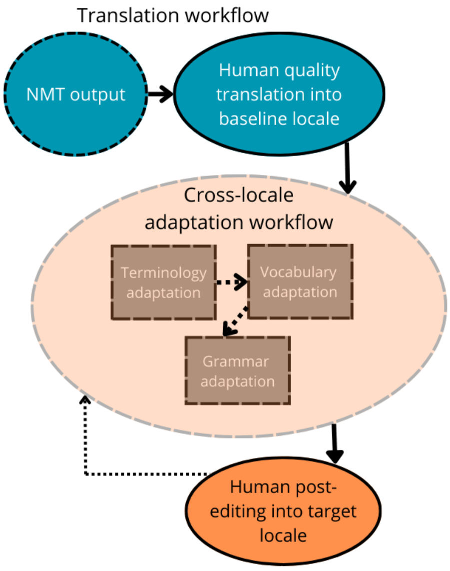

Figure 1: Schema of the proposed workflow. The optional steps are indicated with a discontinuous line.

The proposed workflow for cross-locale adaptation is illustrated in Figure 1. A human-quality translation of the baseline ES variant (with or without prior NMT involvement) goes through a modular cross-locale adaptation process. Depending on the specific types of divergence between the baseline and each target locale, one or more of its components come into play. The adaptation final output is reviewed by human linguists, whose primary focus is ideally on style. However, feedback on grammar or vocabulary can be reintegrated into the system for future use: new lexical items may be added to the vocabulary table with corresponding replacement rules, and grammatical error patterns found can be transformed into fine-tuning examples to improve the model's performance.

Locale	Steps needed	BLEU	PE Distance
MX	1	93.55	4%
AR	1, 2, 3	93.97	4%

Table 3: Adaptation steps needed for each locale transformation and their impact on editing effort. Step 1 corresponds to terminology, Step 2 to vocabulary and Step 3 to grammatical adaptations.

5 Results

To evaluate the results of our experiments, we compared them to the reference translations, also using BLEU and Levenshtein Edit Distance metrics, which let us assess the degree of improvement from the starting point (reflected in Table 1). As shown in Table 3, for ES-AR, the approach demonstrated a reduction in editing effort, with improvements of 9.44 in BLEU scores and 3% in Edit Distance for the chosen workflow. For ES-MX, the improvement is more modest: only 0.86 in BLEU, and 1% in Edit Distance.

To gain a deeper understanding of the results, we asked ES-AR and ES-MX native linguists to review the segments where the adaptation output differed from the reference translation. They were asked to classify each sentence into one of three categories: *acceptable differences* (alternative translations that are equally appropriate for the locale and content type), *minor errors* (slightly inadequate but still intelligible or contextually plausible translations), and *critical errors* (unacceptable mistakes that compromise correctness or clarity in the given context). This additional review and categorization was necessary because the translation metrics used capture deviation from a reference, but do not account for the possibility of multiple valid renderings. Therefore, a lower score does not necessarily indicate that a segment is incorrect or unsuitable.

As Figure 2 shows, from a sample of 1474 segments, out of the 262 ES-LA translations that initially differed from the ES-MX reference, 17% (46) were perfectly adapted to match the ES-MX translations, while 216 were not adapted to exactly match the reference. Among those, only 3% (7) were identified as critical errors, 10% (26) as minor er-

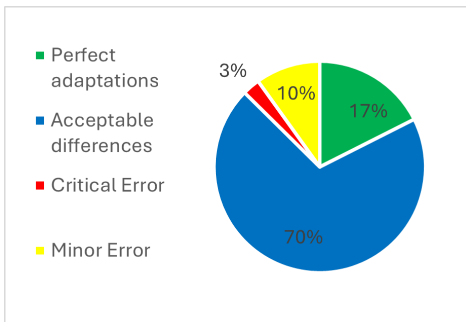

Figure 2: ES-LA segments adapted to ES-MX. The green and blue areas represent the segments that don't need further adaptation, while the red and yellow represent those that do.

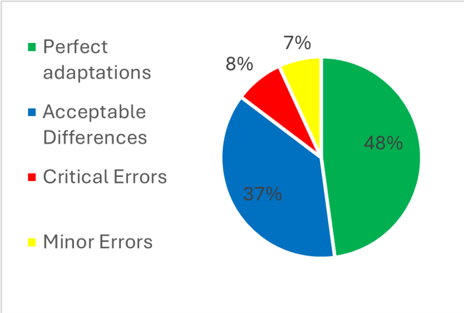

Figure 3: ES-LA segments adapted to ES-AR. The green and blue areas represent the segments that don't need further adaptation, while the red and yellow represent those that do.

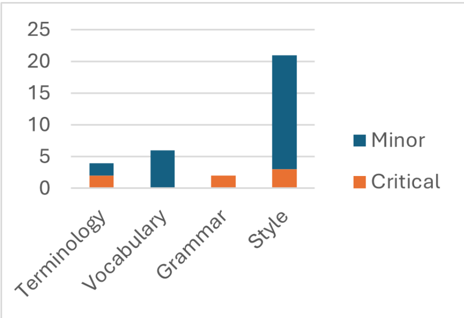

Figure 4: ES-LA into ES-MX error level and typology distribution. Out of the 33 segments with errors, 4 were related to terminology (2 critical, 2 minor), 6 to vocabulary (all minor errors), 2 to grammar (both critical), and 21 to style (3 critical, 18 minor).

rors, and 70% (183) as not requiring further adaptation. As for ES-AR, Figure 3 shows that from a sample of 1567 segments, out of the 625 ES-LA translations that initially differed from the ES-AR reference, 48% (299) were perfectly adapted to match the ES-AR translations, while 326 were not adapted to match the reference. Among those, 8% (49) were identified as critical errors, 7% (43) as minor errors, and 37% (234) as not requiring further adaptation. We also asked the linguists to classify the segments labeled as "errors", both critical and minimal, into the four categories defined in Section 3: terminology, vocabulary, grammar and style. Results for ES-MX and ES-AR can be found in Figure 4 and Figure 5 respectively, and they show that the objective stated in Section 4.4 was achieved: most of the fixes translators would have to perform pertain to the output's style.

In short, 87% of the ES-MX and 85% of the ES-AR automatically adapted segments would be ready for immediate publication, significantly reducing the amount of human post-editing effort involved.

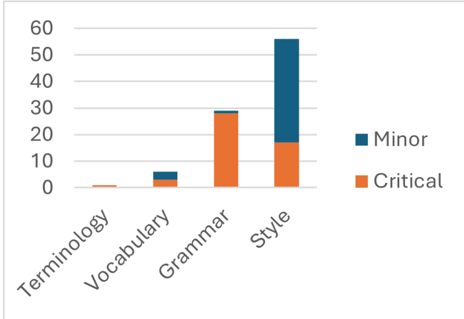

Figure 5: ES-LA into ES-AR error level and typology distribution. Out of the 92 segments with errors, 1 was related to terminology (critical), 6 to vocabulary (3 critical, 3 minor), 29 to grammar (28 critical, 1 minor), and 56 to style (17 critical, 39 minor).

6 Conclusions

In conclusion, this paper has introduced an innovative approach to same-language localization by leveraging the contextual understanding and generative capabilities of LLMs, along with linguistic resources and prompting techniques, to re-imagine the task as more akin to a specific type of post-editing rather than a completely separate process. This method provides a deeper understanding of translation and localization workflows, mitigating the need for developing and maintaining multiple localization strategies and translation models for the different locales of a language, and allowing us to understand the rich and complex relationships between them.

The results demonstrate that this approach is feasible for marketing and product/UI content in Spanish, both for variants that exhibit multiple types of divergences from the chosen baseline locale and for those presenting just one. While not perfect without subsequent human reviewing, these processes can significantly reduce the implicated human post-editing efforts in the more mechanical type of adjustments, allowing linguists and translators to concentrate almost exclusively on the more creative aspects of their work, mainly related to style and brand identity.

Future steps involve expanding the approach to more language pairs, particularly those comprising non-Romance languages, which would present very different challenges. Furthermore, final data collected through this process could be used to fine-tune an LLM, exploring whether style adaptations —which we did not succeed in automating— can be taught to the LLM through demonstration rather than explicit instructions.

References

Paul Bennett. 2002. Teaching contrastive linguistics for MT. In *Proceedings of the 6th EAMT Workshop: Teaching Machine Translation*, Manchester, England. European Association for Machine Translation.

Carl Chapman and Kathryn T. Stolee. 2016. Exploring regular expression usage and context in python. In *Proceedings of the 25th International Symposium on Software Testing and Analysis*, ISSTA 2016, page 282–293, New York, NY, USA. Association for Computing Machinery.

Keiran J. Dunne and Elena S. Dunne. 2011. *Translation and Localization Project Management: The art of the possible*. John Benjamins Publishing, Amsterdam, The Netherlands and Philadelphia, USA.

Ping Ke. 2019. *Contrastive Linguistics*. Peking University linguistics research. Springer, Singapore.

Loïc Lannelongue, Jason Grealey, and Michael Inouye. 2021. Green algorithms: Quantifying the carbon footprint of computation. *Advanced Science*, 8(12):2100707.

Vladimir I Levenshtein. 1966. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10:707.

Antonio Moreno-Sandoval and José Miguel Goñi-Menoyo. 2002. Spanish inflectional morphology in datr. *Journal of Logic, Language and Information*, 11:79–105.

OpenAI. Fine-tuning. fine-tune models for better results and efficiency. https://platform.openai.com/docs/guides/fine-tuning. Accessed: 2025-02-06.

OpenAI. 2023. A survey of techniques for maximizing llm performance. https://www.youtube.com/watch?v=ahnGLM-RC1Y. Accessed: 2025-01-25.

OpenAI. 2024. Gpt-4o mini: advancing cost-efficient intelligence. Technical report, OpenAI.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Maria Carolina Penteado and Fábio Perez. 2023. Evaluating gpt-3.5 and gpt-4 on grammatical error correction for brazilian portuguese. *arXiv preprint arXiv:2306.15788*. Accepted to LatinX in AI (LXAI) Research at ICML 2023.

Rizwan Qureshi, Muhammad Usman Hadi, Qasem Al-Tashi, Abbas Shah, Amgad Muneer, Muhammad Irfan, Anas Zafar, Naveed Akhtar, Mohammed Al-Garadi, Muhammad Shaikh, Syed Hassan, Maged Shoman, Jia Wu, Seyedali Mirjalili, and Mubarak Shah. 2024. Large language models: A comprehensive survey of its applications, challenges, limitations, and future prospects. Accessed: 2025-01-30.

Vikas Raunak, Amr Sharaf, Yiren Wang, Hany Awadalla, and Arul Menezes. 2023. Leveraging GPT-4 for automatic translation post-editing. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12009–12024, Singapore. Association for Computational Linguistics.

Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni. 2020. Green ai. *Commun. ACM*, 63(12):54–63.

Vera Senderowicz. 2024. From “comment allez-vous?” to “comment ça va?”: Leveraging large language models to automate formality adaptation in translation. In *Proceedings of the 16th Conference of the Association for Machine Translation in the Americas (Volume 2: Presentations)*, pages 237–254, Chicago, USA. Association for Machine Translation in the Americas.

Satoru Uchida. 2024. Using early llms for corpus linguistics: Examining chatgpt’s potential and limitations. *Applied Corpus Linguistics*, 4(1):100089.

Celia Uguet, Fred Bane, Mahmoud Aymo, João Torres, Anna Zaretskaya, and Tània Blanch Miró Blanch Miró. 2024. LLMs in post-translation workflows: Comparing performance in post-editing and error analysis. In *Proceedings of the 25th Annual Conference of the European Association for Machine Translation (Volume 1)*, pages 373–386, Sheffield, UK. European Association for Machine Translation (EAMT).

Sara Zanzottera. 2024. Controlling llm translations of invariant elements with rag. https://drive.google.com/file/d/1BvcNbsAGWp25EDpiQ5ljYos3_eneo3wu/view?t=3. Accessed: 2025-02-07.

A Appendix A. Example prompt to generate terminology adaptations when an equivalent term in ES-LA is available

You are a Spanish-speaking linguist from *Mexico/Argentina*. You are instructed to:

1. Read the original English text: $'original - text'$. Can you find any EN terms from the Translation Guide below in it? If you can’t, stop reading the instructions and don’t do anything else. If you do, go on to step 2 below.

2. Read the Spanish translation: $'spanish - translation'$. If $'en - term'$ is translated as $'esLA - term'$ in the Spanish translation, replace it for $'esMX - term'$/$'esAR - term'$. The replacement should be case-insensitive but

should respect the original capitalization of the term in the text.

Follow these general instructions:

A. Watch out! Don't do a "search and replace" type of job. The terms from the Translation Guide might have a different gender, number or capitalization in the text and still be the same. Example 1: if the Translation Guide includes the term "phones", and you find "phone" in the English text, you can consider them a match. Example 2: if the Translation Guide includes the term "callejón" and you find "Callejones" in the Spanish translation, you can consider them a match, even if the word is in plural and capitalized. Be smart about that when you're editing.

B. Morphology matters a lot in Spanish: When you replace Spanish word for another, make sure all articles and adjectives related are adapted accordingly. Don't produce outputs like "El chica" or "Las guapos altas", which are agrammatical in Spanish. The same goes for verbs: when you do replacements, make sure the original conjugation from the text in Spanish is respected.

C. After applying only those changes, return the final version of the translation, without any extra words, explanations, or headers.

Translation Guide:

EN term -> esLA term -> esMX term | esAR term

B Appendix B. Example prompt to generate terminology adaptations when no equivalent term in ES-LA is available

You are a Spanish-speaking linguist from *Mexico/Argentina*. You are instructed to:

1. Read the original English text: $'original - text'$. Can you find any EN terms from the Translation Guide below in it? If you can't, stop reading the instructions and don't do anything else. If you do, go on to step 2 below.

2. Read the Spanish translation: $'spanish - translation'$. Make sure $'en - term'$ is translated as $'esMX - term'|'esAR - term'$ in the Spanish text, and make necessary adjustments if it's not.

Follow these general instructions:

A. Watch out! Don't do a "search and replace" type of job. The terms from the Translation Guide might have a different gender, number or capitalization in the text and still be the same. Example 1: if the Translation Guide includes the term 'phones', and you find "phone" in the English text, you can consider them a match. Example 2: if the Translation Guide includes the term "callejón" and you find "Callejones" in the Spanish translation, you can consider them a match, even if the word is in plural and capitalized. Be smart about that when you're editing.

B. Morphology matters a lot in Spanish: When you replace Spanish word for another, make sure all articles and adjectives related are adapted accordingly. Don't produce outputs like "El chica" or "Las guapos altas", which are agrammatical in Spanish. The same goes for verbs: when you do replacements, make sure the original conjugation from the text in Spanish is respected.

C. After applying only those changes, return the final version of the translation, without any extra words, explanations, or headers.

Translation Guide:

EN term -> No-term -> esMX term | esAR term

C Appendix C. Example prompt to generate terminology adaptations when the Translation Guide includes more than one term

You are a Spanish-speaking linguist from *Mexico/Argentina*. You are instructed to:

1. Read the original English text: $'original - text'$. Can you find any EN terms from the Translation Guide below in it? If you can't, stop reading the instructions and don't do anything else. If you do, go on to step 2 below.

2. Read the Spanish translation: $'spanish - translation'$. Make sure that every EN term is translated as its corresponding esMX term in the Spanish translation, and not as its esLA term. Make the necessary replacements to make that true. The replacement should be case-insensitive but should respect the original capitalization of the term in the text.

Follow these general instructions:

A. Watch out! Don't do a "search and replace" type of job. The terms from the Translation Guide might have a different gender, number or capitalization in the text and still be the same. Example 1: if the Translation Guide includes the term "phones", and you find "phone" in the English text, you can consider them a match. Example 2: if the Translation Guide includes the term "callejón" and you find "Callejones" in the Spanish translation, you can consider them a match, even if the word is in plural and capitalized. Be smart about that when you're editing.

B. Morphology matters a lot in Spanish: When you replace Spanish word for another, make sure all articles and adjectives related are adapted accordingly. Don't produce outputs like "El chica" or "Las guapos altas", which are agrammatical in Spanish. The same goes for verbs: when you do replacements, make sure the original conjugation from the text in Spanish is respected.

C. After applying only those changes, return the final version of the translation, without any extra words, explanations, or headers.

Translation Guide:

EN term -> No-term -> esMX term | esAR term

EN term -> esLA term -> esMX term | esAR term

EN term -> No-term -> esMX term | esAR term

D Appendix D. Example prompt to generate ES-AR vocabulary adaptations

You are a Spanish-speaking linguist from Argentina, specialized in Spanish locale adaptation. Adapt the given Spanish translation according to the following steps:

Approach this task step-by-step, in the specified order, take your time and do not skip steps.

1. Read the Spanish translation carefully: $'spanish - translation'$.

2. Change any future tense verbs to the "ir a" + infinitive form.

3. Change any present perfect form (verb "haber" + past participle) into simple past.

4. Change specific words. Convert:

- "aquí" to "acá",
- "aún" to "todavía",
- "luego" to "después",
- the verb "presionar" into "tocar",
- the verb "permitir" into "dejar",
- the verb "utilizar" into "usar",
- the verb "deber" into the construction "tener que", when applicable, respecting the original conjugation.

After applying the listed changes, make sure the result is still a good translation of $'original - text'$. Then return the final version of the translation. If no changes are applicable, return "No response". Do not add any extra words, explanations, or headers. Do not translate any content into English.

E Appendix E. Example prompt to generate ES-AR grammatical adaptations

You are a Spanish-speaking linguist from Argentina, specialized in Spanish locale adaptation. Adapt the given Spanish translation according to the following steps:

Approach this task step-by-step, in the specified order, take your time and do not skip steps.

1. Read the Spanish translation carefully: $'spanish - translation'$. 2. Transform any second person verbs and pronouns to their Argentine Spanish form using "vos"/"ustedes". 3. After applying the listed changes, make sure the result is still a good translation of $'original - text'$. Then return the final version of the translation.

If no changes are applicable, return "No response". Do not add any extra words, explanations, or headers. Do not translate any content into English.