Code-Switching Grammaticality in LMs

Elizaveta Kuznetsova
University of California, San Diego

Abstract

We investigate how transformer-based language models process bilingual code-switched sentences, evaluating the presence of grammatical knowledge in LMs. Therefore, we create a dataset of minimal pairs contrasting grammatical and ungrammatical code-switched sentences in Spanish-English and English-Spanish directions. Using the XLM-Roberta language model and Microsoft Copilot, we compute sentence-level surprisal scores to evaluate model preferences for grammatical over ungrammatical sentences. Our analysis reveals that models generally capture local agreement phenomena well but exhibit limitations with complex noun phrases and more complex nested structures. These results highlight gaps in current language models' ability to integrate cross-linguistic structural constraints. All project materials are available at https://github.com/Ellieshka/Code-Switching-Grammaticality-in-LMs.

1 Introduction

The process where bilinguals merge two or more languages into one coherent whole Broersma and de Bot (2006) is called code-switching. Code-switching occurs when a speaker alternates between languages within the same discourse, clause, or even within a single word. It is not random or chaotic, but rather it follows linguistic, cognitive, and pragmatic rules. As Bhattacharya and van Schijndel (2024) and Calvillo et al. (2020) note, bilinguals regularly switch languages in natural contexts, and understanding when and why they do so offers valuable insights into how the bilingual mind organizes and processes language. To illustrate the notion of code-switching, the following example was taken from the actual corpus that will be used for analysis:

```
(1) como si fuera a darte un hug
"as if I were going to give you a hug"
```

Here, the speakers engage in conversation in Spanish but insert the English word "hug". While this is a simple single-word insertion, it demonstrates the phenomenon clearly. Such instances raise interesting questions about why speakers switch languages at particular points and whether computational models, like language models, can predict or "get surprised" by these switches.

Despite growing interest in computational modeling of bilingual phenomena, relatively few studies have investigated whether large language models possess sensitivity to the grammatical constraints of code-switching. Evaluations using frameworks such as BLiMP (Warstadt et al., 2020) have demonstrated that models can capture subtle syntactic distinctions in monolingual English, but whether such sensitivity extends to code-switched contexts remains unclear.

The present study examines whether the large language model, Copilot (Microsoft, 2025) in our case, possesses the innate grammatical knowledge about which sentences are grammatical and which ones are not, with emphasis on code-switching sentences (Spanish-English language pair), following BLiMP evaluation framework (Warstadt et al., 2020), which included English-only minimal pairs. We hypothesize that the language model surprisal for the grammatical sentence would be lower than for the ungrammatical one, where accuracy, measuring how often the surprisal for the ungrammatical one is higher, was used to satisfy this condition. To place this investigation within existing research, it is important to consider frameworks that describe how bilingual speakers structure code-switched speech. Understanding these structural patterns provides the theoretical basis for evaluating whether language models capture grammatical constraints in bilingual contexts.

2 Literature Review

Building on these theoretical perspectives, research in bilingual syntax has proposed formal frameworks, such as the Matrix Language Frame model (Myers-Scotton, 1993), to describe the structural constraints governing code-switched speech. According to the MLF framework, bilingual speech consists of a Matrix Language, which provides the grammatical framework, and an Embedded Language, which supplies content morphemes. Because content morphemes carry semantic meaning, the model predicts that they are more likely to be inserted from the Embedded Language, with nouns being the most frequently switched, followed by verbs and other word classes. Nouns are considered "portable" (Myslín and Levy, 2015) because their meaning can map readily across languages.

Language models assign probabilities to sequences of words. Current state-of-the-art transformers have become dominant in neural networks, but evaluation benchmarks like GLUE (Wang et al., 2018) do not fully capture models' knowledge of grammar. Domain-general corpora like CoLA (Warstadt et al., 2019) provide labeled sentences for acceptability judgments and have been used to track model performance. CoLA evaluations require supervised training, limiting conclusions about a model's innate grammatical knowledge. Evaluating LMs directly on minimal pairs avoids this limitation, using sentence probabilities as a confounding factor for acceptability. This motivation underlies the design of BLiMP (Warstadt et al., 2020).

Building on these probabilistic perspectives, recent computational approaches have examined code-switching through the lens of language model surprisal, a measure of word predictability derived from probabilistic models of language. Higher surprisal indicates that a word is less expected given the preceding context, while lower surprisal indicates greater predictability. According to Calvillo et al. (2020), surprisal was found to predict code-switching, i.e., the probability of switching increases with higher-surprisal. On the contrary, when a word is harder to predict, bilinguals may switch to ease production or highlight highly informative content (Myslín and Levy, 2015). Bhattacharya and van Schijndel (2024) show that bilinguals maintain fine-grained syntactic expectations across languages, a finding that aligns with surprisal-based evaluation metrics for

neural language models. Taken together, these theories suggest that code-switching is influenced by a complex interaction of grammatical constraints, lexical accessibility, discourse context, and cognitive effort.

These theoretical and computational perspectives motivate the present study, which examines whether Copilot (Microsoft, 2025) has necessary grammatical knowledge to differentiate between grammatical and ungrammatical code-switching in Spanish-English sentences. By combining linguistic theory with surprisal, we aim to investigate whether the model can distinguish grammatical from ungrammatical code-switched sentences.

3 Methods

3.1 Phenomena Scheme

The linguistic phenomena examined in this study were defined in consultation with Anne Beatty-Martínez, a code-switching scholar and Principal Investigator of the Bilingualism in Context Lab at the University of California, San Diego. The following eight phenomena were selected to illustrate Spanish–English code-switching:

- Phenomenon 1: Simple Noun Phrases (Simple NPs)
  This phenomenon targets sentences containing simple noun phrases consisting of a single noun (e.g., the book, el libro).

- Phenomenon 2: Complex Noun Phrases (Adjective-Noun)
  This category includes noun phrases composed of an adjective and a noun (e.g., the red book, el libro rojo).

- Phenomenon 3: Complex Compound Noun Phrases
  This phenomenon examines noun phrases containing multiple modifiers or compound structures (e.g., the small red book on the table).

- Phenomenon 4: Adjective-Noun Word Order
  This category tests whether adjective–noun ordering conforms to the syntactic rules of the relevant language (e.g., red book vs. libro rojo).

- Phenomenon 5: Auxiliary Verb Switch
  This phenomenon focuses on sentences in which the auxiliary verb switches languages

(e.g., has comido → has eaten). Tense, aspect, and agreement are expected to remain grammatically well-formed.

- Phenomenon 6: Copula-Predicate Switch
This category includes sentences where the copula (e.g., is, está) appears in one language and the predicate in another (e.g., is feliz).

- Phenomenon 7: Subject-Verb Agreement
This category targets sentences requiring subject–verb agreement in person and number (e.g., he eats, él come).

- Phenomenon 8: Verb-Direct Object Agreement
This phenomenon examines sentences in which the main verb switches languages while maintaining appropriate agreement with the direct object (e.g., number or gender).

3.2 Language Model

The large language model Copilot (Microsoft, 2025) is a generative language model built on the GPT architecture (OpenAI, 2023). Our experiments use the version integrated into GitHub Copilot, which is based on GPT-3.5, containing approximately 175B parameters. The model is pretrained on a mixture of publicly available code and text data, including repositories from GitHub and other large-scale web text corpora. Exact training data do not seem to be publicly disclosed, but estimates suggest the model was exposed to billions of tokens of natural language and code.

3.3 Script

We first converted our stimuli into a format compatible with the minicons evaluation framework. Using pandas, the CSV file containing all minimal pairs was reshaped from wide to long format. Each sentence was assigned a unique label combining the phenomenon, pair ID, and condition (grammatical/ungrammatical; switch direction), and only the columns required by minicons (label and sentence) remained.

We computed surprisal scores for each sentence using the minicons library (**?**) with the xlm-roberta-base language model. The model processed sentences in batches of 50 to balance memory usage and efficiency. Surprisal was calculated as the negative sum of the log-probabilities of each token in the sequence. These scores were then appended to the original DataFrame, resulting in a complete mapping of sentences to surprisal values for subsequent analysis.

To evaluate a model's grammaticality knowledge, we checked whether the surprisal of the grammatical sentence was lower than that of the ungrammatical counterpart or each minimal pair. This produced a binary preference metric, which was then averaged across phenomena to calculate accuracy per phenomenon.

Finally, we visualized the surprisal distributions using seaborn. Boxplots were generated to compare surprisal by code-switch direction and grammaticality, showing us how surprisal patterns varied across structural conditions. For details on the code, please see the repository link.

3.4 Stimulus Design and Generation

The pairs were designed to be full sentences, rather than short phrases, following BLiMP's emphasis on sentence-level evaluation (Warstadt et al., 2019). All datasets were generated artificially, allowing the isolation of a single contrast within each minimal pair (Ettinger et al., 2018). For each phenomenon, sentences were generated using Copilot (Microsoft, 2025) guided by conditional prompts (see Appendix A) specifying the relevant syntactic configuration, code-switching direction, and grammatical manipulation. Prompts were designed to produce paired grammatical and ungrammatical sentences that were matched in length and lexical content, differing only in the targeted violation.

Rather than sampling from a predefined vocabulary or relying on explicit lexical annotation, grammatical well-formedness and selectional compatibility were enforced through prompt constraints and subsequent manual inspection during data restructuring. This procedure allowed for controlled stimulus generation while maintaining naturalness and minimizing unintended variation across minimal pairs. Sentences were generated in batches of 50, after which the model prompted the user to continue. Upon each confirmation, an additional batch of 50 sentences was produced. This process continued until stimulus generation was complete. For minimal pair samples, please see Appendix B.

3.5 Data Organization and Cleaning

All generated sentences were first compiled into a Google Sheets file for restructuring, labeling, and data cleaning. During this stage, sentences were manually inspected to correct formatting inconsistencies, remove duplicates, and ensure that gram-

matical and ungrammatical versions differed only in the targeted phenomenon. Each sentence was annotated with metadata including phenomenon type, pair ID, code-switch direction, and grammaticality label.

Once initial cleaning and annotation were complete, each pair of phenomena (1–2, 3–4, 5–6, 7–8) was organized into a separate Google Sheet to facilitate systematic validation. These sheets were then shared with human validators, who provided acceptability judgments to confirm that the grammatical labels aligned with native speaker intuitions. Sentences with inconsistent judgments or unclear acceptability were reviewed and revised or removed prior to inclusion in the final dataset.

3.6 Human Validation

Only phenomena 5–6 were validated by a native speaker of Spanish from Spain. The validator identified several issues, including dialectal word choice, stylistic inconsistencies, typographical errors, duplicate items within the same phenomenon, and other discrepancies.

Overall, 15 out of 400 minimal pairs raised concerns. These items were not removed at this stage, as the remaining phenomena were not validated due to limited human resources. Removing these pairs would result in an unequal distribution of minimal pairs across the four phenomenon sets, with phenomena 5–6 containing fewer items than the others.

3.7 Labeling Scheme

Each sentence was assigned a composite label encoding the phenomenon, sentence index, language switch direction, and grammaticality. For example:

`SimpleNPs_1_spa_eng_grammatical`

Here, `SimpleNPs` indicates the phenomenon, `1` the sentence number within that phenomenon, `spa_eng` the direction of code-switching, and grammatical the acceptability status. The dataset was converted to a long format to facilitate automated analysis and compatibility with surprisal computation and statistical aggregation.

3.8 Surprisal Computation

Surprisal was evaluated at the utterance level using Minicons (**?**), computed as the negative log-probability assigned by the language model to a complete sentence. For each minimal pair, an accuracy score was derived from the following hypothesis:

```
surprisal(gram) < surprisal(ungram)
```

The language model is assumed to prefer the grammatical sentence.

For each sentence, total surprisal was calculated as the sum of token-wise surprisals. For each minimal pair, the surprisal values of the grammatical and ungrammatical sentences were compared to determine model preference. Surprisal was computed over the full sentence, rather than at a specific token position, to capture global sentence well-formedness.

Accuracy scores were averaged per phenomenon to assess model performance. Scores closer to 1.0 indicate higher accuracy, while scores closer to 0.0 indicate poorer performance.

If the above hypothesis is true, then the language model chooses a grammatical sentence, therefore, it makes an accurate judgement. This approach follows the evaluation paradigm used in BLiMP (Warstadt et al., 2020), where higher probability (lower surprisal) for grammatical sentences is interpreted as evidence of grammatical sensitivity.

4 Results

4.1 Performance by Phenomenon

As we can see from Table 1, The model performed best on SimpleNPs, AuxVerb, Subject–Verb Agreement, and Copula–Predicate phenomena. Moderate performance was observed for ComplexNPs, CompoundNPs, and Verb–Direct Object Agreement. The weakest performance was found for Adjective Order.

Table 1: Accuracy per Phenomena Switch Direction

Phenomenon	Switch Direction	Accuracy
Adj-Noun	Eng-Spa/Spa-Eng	0.425
Aux Verb	Eng-Spa/Spa-Eng	0.995
Simple NPs	Eng-Spa/Spa-Eng	1.000
Complex NPs	Eng-Spa/Spa-Eng	0.660
Compound NPs	Eng-Spa/Spa-Eng	0.660
Copula Pred	Eng-Spa/Spa-Eng	0.875
Subj-Verb Agr	Eng-Spa/Spa-Eng	0.940
Verb-DO Agr	Eng-Spa/Spa-Eng	0.670

4.2 Performance by Language Direction

No differences in accuracy scores Table 1 were observed between the two code-switching directions (Spanish-English vs. English-Spanish). One
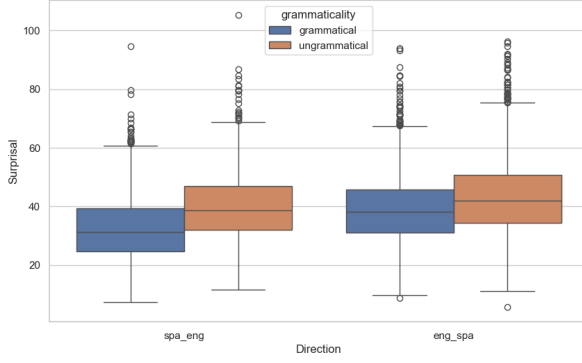
Figure 1: Surprisal Distributions by Directionality

possible explanation is that the stimuli were generated by a language model and were not fully validated by human annotators. As a result, the model may have produced structurally mirrored sentence pairs, yielding identical surprisal values across both switching directions.

4.3 Distributional Analysis

Figure 1 shows the distribution of utterance-level surprisal values by grammaticality and code-switching direction. Across both Spanish→English and English→Spanish switches, ungrammatical sentences exhibit consistently higher median surprisal than the grammatical ones, despite identical binary accuracy scores across directions. At the same time, the distributions show substantial overlap, as well as greater variance and heavier upper tails for ungrammatical sentences. These patterns indicate that, while the model often assigns higher surprisal to ungrammatical inputs, the strength of this preference varies across items. Consequently, distributional analyses show that models respond to grammaticality in varying degrees, which is hidden if we only look at simple accuracy scores. This gives a more detailed picture of model behavior.

5 Discussion

The results reveal varied findings in model performance across linguistic domains. Consistent with the MLF model (Myers-Scotton, 1993), nouns are handled well, but complex structures reveal limitations in the model's ability to integrate hierarchical syntactic constraints across languages In addition, Auxiliary Verb Switches, Copula-Predicate Switches, and Subject-Verb Agreement show a very high performance. This pattern suggests that the model robustly encodes basic morphological structure in code-switching contexts.

In contrast, phenomena related to word order

or more complex internal noun phrase structure present greater difficulty. Adjective-Noun Word Order shows the weakest performance, indicating that the model struggles to accurately identify word order constraints within the code-switching realm. Similarly, Complex and Compound Noun Phrases and Verb-Direct Object Agreement demonstrate only moderate accuracy, suggesting limitations in representing relational dependencies across languages.

These results parallel findings from BLiMP (Warstadt et al., 2020), where the model performs strongest on local agreement phenomena and weakest on domains requiring abstract structural generalizations. However, high performance on a given phenomenon should not be interpreted as evidence of fully human-like grammatical knowledge, but rather as sensitivity to particular surface-level regularities present in the evaluation data.

6 Limitations

Several limitations should be noted. First, unlike Warstadt et al. (2020) study, we didn't rely on a predefined vocabulary or lexical annotation, which could contribute to more robust or even opposite results. Second, only a subset of phenomena were validated by humans due to resource constraints, which may have influenced the switch direction effects in model's accuracy scores. Thirdly, we used only one model for computing surprisal, whereas a comparison of several large language models could show more robust or completely different results. Verb-Direct Object Agreement Phenomenon may not be entirely valid for English. All the number agreement pairs conform fully to standard English grammar in English-dominant sentences, which limits the phenomenon's ability to test the model's sensitivity to agreement violations. Moreover, the model produced non-variate pairs, some of them were duplicated under the same phenomena, and some of them just used the same verb for all the pairs, which might have skewed the results significantly. Unlike BLiMP, the present study does not investigate the effect of sentence length, model confidence, or training size on performance. Because the evaluation focuses on a single pretrained model and full-sentence surprisal comparisons, we leave a more fine-grained analysis of shallow predictors of performance, such as length or lexical frequency for future work. Nevertheless, examining surprisal distributions in addition to binary accuracy pro-

vides initial evidence that model preferences are graded rather than categorical. Finally, the analysis focuses on a single model configuration, limiting generalizability across architectures. Despite these limitations, the study provides a controlled and scalable framework for evaluating grammatical knowledge in code-switching contexts.

References

Tanmoy Bhattacharya and Marten van Schijndel. 2024. Large language models implicitly learn where humans code-switch. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)*, pages 1–31.

Mirjam Broersma and Kees de Bot. 2006. Triggering in code-switching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, 9(1):1–13.

Eduardo Calvillo, Jingyang Zhang, and Debra Titone. 2020. Code-switching is predictable: A computational and empirical study of surprisal in bilingual sentence processing. *Bilingualism: Language and Cognition*, 23(3):681–692.

Allyson Ettinger, Ahmed Elgohary, Colin Phillips, and Philip Resnik. 2018. Assessing composition in sentence vector representations. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1790–1801.

Microsoft. 2025. Copilot. https://copilot.microsoft.com/. Large language model.

Carol Myers-Scotton. 1993. *Duelling languages: Grammatical structure in code-switching*. Clarendon Press, Oxford, UK.

Mark Myslín and Roger Levy. 2015. Code-switching and predictability of meaning in discourse. *Language*, 91(4):871–905.

OpenAI. 2023. Chatgpt: Optimizing language models for dialogue. https://openai.com/chatgpt.

Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R. Bowman. 2018. Glue: A multi-task benchmark and analysis platform for natural language understanding. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 353–355.

Alex Warstadt, Alicia Parrish, Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, and Samuel R. Bowman. 2020. Blimp: The benchmark of linguistic minimal pairs for english. *Transactions of the Association for Computational Linguistics*, 8:377–392.

Alex Warstadt, Amanpreet Singh, and Samuel R. Bowman. 2019. Neural network acceptability judgments. *Transactions of the Association for Computational Linguistics*, 7:625–641.

A Stimuli Generation Prompt

A.1 MASTER PROMPT

Task: Generate code-switching minimal pairs for a specified phenomenon. Each minimal pair consists of a grammatical and an ungrammatical sentence for both switch directions (Spanish-English and English-Spanish) that differ only in the manipulated switch defined for that phenomenon.

A.2 GLOBAL CONSTRAINTS

Produce full sentences, not fragments. Sentences must sound natural except for the manipulated switch. For each minimal pair:

- Keep all lexical material identical except for the switch.

- Keep the switch point identical across grammatical and ungrammatical versions.

- Do not introduce or remove any words other than what the switch requires.

Embed the target construction inside a complete sentence with identical surrounding context across the pair. Use masculine Spanish nouns wherever relevant. Respect language-specific word order unless the phenomenon explicitly targets word order. Generate 200 minimal pairs per phenomenon: 50 examples for each of the four pair types.

A.3 GENERAL GENERATION INSTRUCTIONS

Select appropriate lexical items that satisfy the phenomenon-specific constraints.
For each example, generate a grammatical sentence first, then its ungrammatical counterpart.
Ensure that the only difference between the two is the specified switch.
Label each sentence with its corresponding Type.
Do not include explanations, comments, or numbering.

A.4 PHENOMENON-SPECIFIC PARAMETERS

A.4.1 PHENOMENON 1: SIMPLE NPs (Determiner–Noun)

Manipulated switch: determiner–noun order and language.

Lexical constraints: Use only English nouns whose Spanish translation is masculine.
Four pair types:

-
- Spa→Eng grammatical: el + English noun
- Spa→Eng ungrammatical: English noun + el
- Eng→Spa grammatical: the + Spanish noun
- Eng→Spa ungrammatical: Spanish noun + the

Additional constraint: Only the determiner–noun configuration may change.

A.4.2 PHENOMENON 2: COMPLEX NPs (Adjective–Noun)

Manipulated switch: adjective–noun order and language.
Lexical constraints: Nouns must have masculine Spanish equivalents.
Four pair types:

- Spa→Eng grammatical: el + English adjective + English noun
- Spa→Eng ungrammatical: el + English noun + English adjective
- Eng→Spa grammatical: the + English noun + Spanish adjective
- Eng→Spa ungrammatical: the + Spanish adjective + English noun

A.4.3 PHENOMENON 3: COMPLEX COMPOUND NPs

Manipulated switch: compound structure vs. Spanish de construction.
Lexical constraints: Use nouns with masculine Spanish equivalents.
Four pair types:

- Spa→Eng grammatical: el + English compound noun (e.g., fiction book)
- Spa→Eng ungrammatical: el + English noun + de ficción
- Eng→Spa grammatical: the + English noun + de ficción
- Eng→Spa ungrammatical: the + English compound noun

A.4.4 PHENOMENON 4: ADJECTIVE–NOUN WORD ORDER

Manipulated switch: prenominal vs. postnominal adjective position across languages.
Lexical constraints: Use masculine Spanish nouns.
Four pair types:

- Spa→Eng grammatical: el + English adjective + Spanish noun
- Spa→Eng ungrammatical: el + Spanish noun + English adjective
- Eng→Spa grammatical: the + English noun + Spanish adjective
- Eng→Spa ungrammatical: the + Spanish adjective + English noun

A.4.5 PHENOMENON 5: AUXILIARY VERB SWITCH

Manipulated switch: auxiliary–participle compatibility.
Lexical constraints: Use plausible activity, motion, or perception verbs.
Four pair types:

- Spa→Eng grammatical: está + English -ing form
- Spa→Eng ungrammatical: está + English bare verb
- Eng→Spa grammatical: is + Spanish -ando/-endo form
- Eng→Spa ungrammatical: is + Spanish infinitive

A.4.6 PHENOMENON 6: COPULA–PREDICATE SWITCH

Manipulated switch: copula compatibility with predicate type.
Lexical constraints: Use adjectives denoting physical or mental states. Use Spanish
infinitives that clearly cannot act as adjectives.
Four pair types:

- Spa→Eng grammatical: está + English adjective
- Spa→Eng ungrammatical: está + English bare verb
- Eng→Spa grammatical: is + Spanish adjective
- Eng→Spa ungrammatical: is + Spanish infinitive

A.4.7 PHENOMENON 7: VERB–DIRECT OBJECT AGREEMENT

Manipulated switch: number agreement between verb and direct object.

Lexical constraints: Use countable nouns only.

Four pair types:

- Spa→Eng grammatical: tiene + English plural noun

- Spa→Eng ungrammatical: tiene + English singular noun

- Eng→Spa grammatical: has + Spanish plural noun

- Eng→Spa ungrammatical: has + Spanish singular noun

A.4.8 PHENOMENON 8: SUBJECT–VERB AGREEMENT

Manipulated switch: subject–verb agreement across languages.

Lexical constraints: Subject NP must remain identical across pair members.

Four pair types:

- Spa→Eng grammatical: La figura + has

- Spa→Eng ungrammatical: La figura + have

- Eng→Spa grammatical: The figure + tiene

- Eng→Spa ungrammatical: The figure + tienen

A.5 FINAL INSTRUCTION

Begin generation for the selected phenomenon. Produce exactly 200 sentences (50 per type) following all global and phenomenon-specific constraints.

B Sample Minimal Pairs

Table 2: Sample Minimal Pairs

Phenomenon	N	Spa-Eng Grammatical	Spa-Eng Ungrammatical	Eng-Spa Grammatical	Eng-Spa Ungrammatical
Simple Noun Phrases	1	María dejó el **libro** sobre la mesa antes de la reunión.	María dejó la **libro** sobre la mesa antes de la reunión.	She found the **book** on the table during the meeting yesterday.	She found the **book** on the table during the yesterday meeting.
Complex Noun Phrases	2	El vecino lavó **el carro que estaba muy sucio** en la tarde.	El vecino lavó **carro que estaba muy sucio** en la tarde.	They parked the **old red car** in front of the school before the storm.	They parked in front of the school the **old red car** before the storm.
Complex Noun Phrases	52	La lecto**ra del libro con la portada azul** no asistió a la reunión.	La lecto**ra libro con la portada azul** no asistió a la reunión.	He shelved the **heavy book with the torn cover** after the last page.	He shelved the **heavy book with torn the cover** after the last page.
Adjective-Noun Order	114	Coloca el **libro azul** en la mesa.	Coloca el **azul libro** en la mesa.	Put the **blue book** on the table.	Put the **book blue** on the table.
Auxiliary Verb	65	El niño **ha comido** su cena.	El niño **ha cena** su comido.	He **has eaten** his dinner.	He **eaten has** his dinner.
Copula	41	El obrero **está cansado** por el duro trabajo.	El obrero **cansado está** por el duro trabajo.	The worker **is tired** from the hard work.	The worker **tired is** from the hard work.
Subject-Verb Agreement	170	El iglú **has been** reconstruido.	El iglú **have been** reconstruido.	The ice**house has been** rebuilt.	The ice**house have been** rebuilt.
Verb-Object Agreement	200	El estuche tiene **dos** lápices de colores.	El estuche tiene **dos** lápiz de colores.	The pencil case has **two** colored pencils.	The pencil case has **two** colored pencil.