Training, Education, and Learning Sciences, Vol. 155, 2024, 216–222
https://doi.org/10.54941/ahfe1005406

French as a Second Language (L2) and AI: Deep Learning Models to the Rescue of Object Clitics

Adel Jebali

Concordia University, Montreal, Quebec, Canada

ABSTRACT

French, like other Romance languages, has object clitics that pose a challenge for non-native speakers due to their unique characteristics and positioning. Learners often make errors or use avoidance strategies. We believe deep learning could assist these learners. We trained a model to identify sentences with errors in the use of clitic object pronouns using a substantial amount of training data. We adapted a pre-trained FlauBERT model for a grammaticality judgment task and fine-tuned it on a dataset of sequences annotated as correct or containing errors. The model achieved an overall F-score of 0.93 on a second evaluation corpus. We further fine-tuned the model on a corpus of examples of errors related to the use of these clitics, achieving an evaluation F-score of 0.95. Both models can be integrated into an automatic correction system to help French L2 learners avoid errors related to the use of clitic object pronouns. The system checks for errors in the use of the object clitic and provides feedback based on the error type. An additional generative module was later added, fine-tuned on another corpus and based on mBARThez. Its aim is to reformulate the learner's input and suggest a correction.

Keywords: French second language, Clitics, Deep learning, BERT, BART

INTRODUCTION

Object clitics (OCs) in French are personal or oblique pronouns that are commonly used in sentences such as example (1b):

(1) a. Sarah a écrit ce roman policier.
b. Sarah **l'**a écrit.

In sentence (1b), the OC pronoun "*le*" (which is elided to *l'*) is a syntactic object of the verb and attaches preverbally to it, despite being a prosodically weak element. This is different from the behavior of equivalent pronouns in English, such as in example (2b), where the pronoun "*it*" is placed to the right of the verb in the same position as the nominal phrase (NP) it replaces, "*this detective novel*".

(2) a. Sarah wrote this detective novel.
b. Sarah wrote **it**.

The peculiar behavior of OCs in French, as well as other unique features of these elements, often causes confusion among French second language (L2)

learners. As a result, these learners use various strategies to compensate for their lack of mastery of these forms. Some authors, including (Wust, 2009), have noted the issue of omission and avoidance, while others, such as (Jebali, 2018), have highlighted strategies like NP repetition. Learners also commonly make errors in the placement of OCs relative to the verb and auxiliary, and in false agreements in gender, number, or person with the antecedent. Additionally, grammatical case errors, discussed in (Emirkanian, Redmond and Jebali, 2021), are also prevalent, such as using the accusative instead of the dative and vice versa.

All these strategies and errors are well-represented in the data used to fine-tune two deep learning models aimed at classifying French L2 learners' productions into two categories, and at classifying their errors relative to OCs' use, as explained in the following sections.

DATA

While we were conducting this research and up to the present day, there is no existing corpus or dataset that allows us to fine-tune deep learning models according to our objectives. Therefore, we initially had to create two datasets:

- The first one is primarily composed of authentic mistakes made by learners during a previous experiment conducted by the researcher (Jebali, 2018). This dataset also includes some fabricated sentences that demonstrate common mistakes missing in our initial corpus. In total, the corpus consists of 2636 examples of various mistakes in the use of OCs. This dataset also includes perfectly correct sentences regarding OC use in French, sourced from various websites. To maintain balance, the number of these correct examples was also 2636, bringing the total to 5272 examples.
- The second dataset consists solely of sentences containing mistakes in the use of OCs. We identified four types of mistakes and used examples from the first dataset, supplemented with additional examples to create four balanced categories, each containing 1734 examples, for a total of 6936. More details about these categories will follow.

The First Dataset

This first dataset is geared towards a binary text classification task. So, in this dataset we gathered text sequences that fall into two classes:

1. Correct
2. Erroneous

This classification is done according to the French language rules regarding the use and placement of OCs. Thus, while example (1b) falls into the first class, example (3) below falls into the second one, where the OC *le* is placed to the right of the past participle *écrit* instead of the pre-auxiliary position.

(3) Sarah a écrit **le**.

The Second Dataset

This dataset is geared towards a 4 categories classification task. So, we used examples from corpus 1, but we added much more examples to have enough data that fall into 4 categories of errors:

1. Cat. 1: Agreement errors (in gender, number or person) regarding the antecedent. In example (4), below, the learner used the OC *les* (plural) instead of the OC *le* (l'), which is singular (like the antecedent *papa*).
2. Cat. 2: Placement errors. The OC is otherwise OK, but it is placed in a non-canonical position, such as in example (3).
3. Cat. 3: Strong pronouns instead of OCs or vice versa. In example (5), for instance, the learner used the strong pronoun *eux* instead of the OC *les*.
4. Cat. 4: Semantic and morphosyntactic errors. This is a mix of 4 types of errors: Case (Accusative instead of Dative, for example), Missing and nul object, doubling errors and confusion between *y*, *en*, le and *ça*. The reason we put these errors together is that they all imply a missanalysis of the argument structure. (6) illustrates an example of a nul object:
   (4) Papa a contracté une maladie qui **les** a rendu très faible.
   (5) Mon chien adore les enfants. Pourtant il aboit quand il voit **eux**.
   (6) Est-ce que tu vas à l'école? Oui, je vais.

MODELS FINE-TUNING & EVALUATION

FlauBERT

FlauBERT (Le et al., 2020) is a language model developed by the French National Centre for Scientific Research (CNRS). Serving as the Francophone counterpart to CamemBERT (Martin et al., 2020), both models trace their roots back to BERT (Devlin et al., 2019).

With a robust architecture comprising 373 million parameters, FlauBERT is specifically tailored for the task of grammaticality judgment and text classification. The model's effectiveness is achieved through a fine-tuning technique using the Hugging Face Transformers library. The process involves a two-step approach: firstly, the transfer learning of a pre-trained model into a binary classifier to discern between correct and incorrect sentences regarding OCs' use. The second fine-tuning step focuses on distinguishing among the four types of errors, demonstrating FlauBERT's *understanding* of grammatical nuances.

Fine-Tuning

The two fine-tuning steps were done on a single consumer GPU with a batch size of 32 and 20 epochs with an early stopping at epoch 9 for the first model and at epoch 11 for the second. The first model achieved a validation accuracy of 0.97 and the second 0.99. Figures 1 and 2 show training loss and validation loss values for the two fine-tuned models:

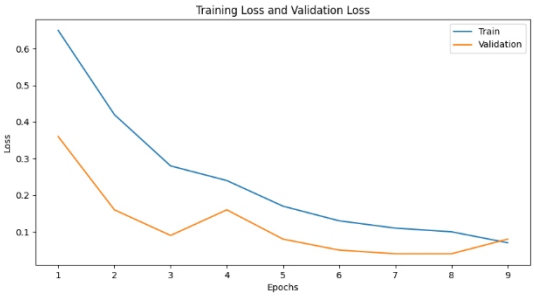

Figure 1: Training and validation loss (model 1).

Figure 2: Training and validation loss (model 2).

Evaluation

During the fine-tuning process, the two datasets were divided into two subsets each, with 90% allocated for training and 10% for evaluation. Following the fine-tuning phase, the models undergo testing using two additional datasets. Models 1 was evaluated on 2282 examples, while Model 2 was tested on 833 examples, all of which have never been encountered by the models before. This approach ensures a robust assessment of the models' generalization capabilities.

These evaluations led to the following results: F-score for model 1 is 0.93 and 0.95 for the second. Confusion matrices are shown in Figures 3 and 4:

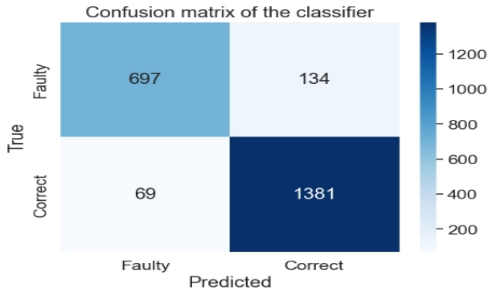

Figure 3: Model 1 eval.

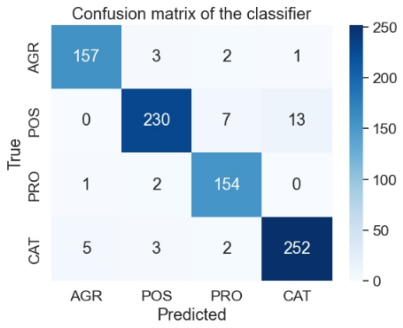

Figure 4: Model 2 eval.

To gauge its performance, comparisons are drawn between model 1 and two benchmark models: Antidote 11 and GPT-3.5 (ChatGPT). These comparisons provide insights into how well the FlauBERT-based model performs in relation to established language tools and advanced large language models. Model 2 cannot be compared to any other tool as this error classification is novel. Table 1 presents this comparison:

Table 1. Comparison results between Model 1, Antidote 11 and GPT 3.5.

	Model 1	Antidote 11	GPT 3.5
Precision	.91	.84	.86
Recall	.84	.38	.54
F-score	.87	.52	.66

Model 1 stands out as notably more robust than both Antidote 11 and GPT 3.5 when it comes to distinguishing between correct and erroneous sentences concerning object clitics (OCs). Conversely, Model 2 exhibits considerable strength in effectively identifying errors made by learners in relation to these OCs with an F1-score of .95, a precision of .95, and a recall of .96 with better results in identifying error types 1 and 3.

Our Models' Potential in Education

These models are fine-tuned in the express purpose of making them useful to French L2 learners. This is achieved by doing the following:

- The integration of these models into a complete system that enables learners to distinguish between correct texts and those containing errors in the use of OCs. It also allows for the identification of the exact type of error, and if the learner wishes, the system can also rephrase the text to eliminate the specific error.
- The feedback provided is tailored to the specific type of error made, ensuring a targeted and effective learning experience.
- The system also highlights and suggests corrections for repetitions in noun phrases (NP) or adjectival phrases (AdjP).

- The system includes a lexical spell-checking module as an initial phase, as spelling corrections enhance the predictions made by our models.
- This system encourages autonomy, empowering learners to take control of their language acquisition journey.
- The rephrase module is based on mBARThez (Eddine et al., 2021) is not chat-like. One of the advantages of this generative module is that it doesn't take the learner's place in telling them what to think. Instead, it operates based on the learner's expressions and ideas, offering linguistic improvements and sometimes even stylistic enhancements.

CONCLUSION

In this paper, we presented a cost-effective system developed by a single individual that will be available for free and will be a valuable tool for French second-language learners. While the system currently has limitations, as it does not address various errors (other than those concerning OCs), some aspects are already covered, including simple agreements, the placement of complements before verbs, negation, punctuation, and more. Looking ahead, the project aims to expand its scope by modeling the interlanguage of French second-language learners, extending beyond just cliticization. This initiative reflects a commitment to enhancing language learning experiences and holds promise for further advancements in supporting learners of French as a second language.

ACKNOWLEDGMENT

This work is made possible through a grant from the Social Sciences and Humanities Research Council (SSHRC) of Canada under the Knowledge Development program. The fine-tuning calculations were performed on the servers of Canada's Digital Alliance, primarily utilizing Niagara and Mist.

REFERENCES

Affes A., I. Biskri and A. Jebali. (2022). French Object Clitics in the Interlanguage: A Linguistic Description and a Formal Analysis in the ACCG Framework, in N. T. Nguyen et al. (Eds.): ICCCI 2022, LNAI 13501, pp. 220–231.

Cheng Y. and M. Duan. (2020). Chinese Grammatical Error Detection Based on BERT Model, Proceedings of the 6th Workshop on Natural Language Processing Techniques for Educational Applications, pp. 108–113.

Devlin J., M.-W. Chang, K. Lee and K. Toutanova. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 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).

Emirkanian E., L. Redmond and A. Jebali. (2021). Maîtrise des clitiques datifs dans les structures bitransitives en français L2 par des apprenants anglophones : Influence de la structure argumentale de la L1, Canadian Journal of Applied Linguistics, Volume 24, Number 2.

Guerdoux G., T. Tiffet and C. Bousquet. (2022). Inference Time of a CamemBERT Deep Learning Model for Sentiment Analysis of COVID Vaccines on Twitter, in J. Mantas et al. (Eds.): Advances in Informatics, Management and Technology in Healthcare.

Jebali A. (2018). Anxiété langagière, communication médiée par les technologies et élicitation des clitiques objets du français L2, Alsic 21.
Kamal Eddine M., A. J.-P. Tixier, and M. Vazirgiannis. (2021). BARThez: A Skilled Pretrained French Sequence-to-Sequence Model, arXiv:2010.12321 [cs. CL].
Liu Y. et al. (2019) RoBERTa: A Robustly Optimized BERT Pretraining Approach, ArXiv, abs/1907.11692.
Martin L. et al. (2020). CamemBERT: A Tasty French Language Model, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 7203–7219.
Noreskal L., I. Eshkol-Taravella and M. Desmets. (2021). Erroneous Coordinated Sentences Detection in French Students' Writings, in K. Wojtkiewicz et al. (Eds.): ICCCI 2021, CCIS 1463, pp. 586–596.
Vaswani A. et al. (2017). Attention Is All You Need, NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, December.
Wust V. (2009). A la recherche des clitiques perdus: The dictogloss as a measure of the comprehension of y and en by L2 learners of French, The Canadian Modern Language Review, Volume 65, Number 3.