Actes de CORIA-TALN 2026
Nantes, France, 29 juin au 3 juillet 2026

SyntaxGym pour le français: Ressource, annotation et évaluation de modèles de langues français et multilingues

Tatiana Bladier Henri-José Deulofeu Alexis Nasr
Aix Marseille Université, CNRS, LIS, UMR 7020, Marseille, France
tatiana.bladier@lis-lab.fr, jose.deulofeu@univ-amu.fr,
alexis.nasr@lis-lab.fr

RÉSUMÉ

Malgré les progrès récents des grands modèles de langues, leur compétence syntaxique reste peu étudiée, en particulier pour les langues autres que l'anglais. Alors que des ressources telles que BLiMP ou SyntaxGym ont permis d'étudier le comportement des modèles de langue sur certaines structures syntaxiques de l'anglais et de l'expagnol, il n'existe rien de comparable pour le français. Pour combler cette lacune, nous présentons SyntaxGymFR, un corpus d'évaluation soigneusement constitué et validé manuellement pour évaluer les capacités syntaxiques des modèles français et multilingues. SyntaxGymFR se compose de paires minimales de phrases ciblant plusieurs phénomènes syntaxiques du français. Nous décrivons la méthodologie d'annotation, la sélection des constructions linguistiques et les procédures de validation utilisées afin de garantir la couverture des phénomènes syntaxiques. De plus, nous présentons des résultats expérimentaux obtenus avec seize modèles français et multilingues, analysant leur sensibilité aux contrastes grammaticaux. Nos résultats fournissent de nouvelles perspectives sur les capacités de généralisation syntaxique des modèles français et établissent SyntaxGymFR comme une ressource pour les recherches futures sur l'évaluation de la compétence syntaxique du français.

ABSTRACT

SyntaxGym for French : Resource, Annotation, and Evaluation of French and Multilingual LLMs

Despite recent advances in large language models (LLMs), their syntactic competence remains insufficiently characterized, especially for languages other than English. While benchmarks such as BLiMP and SyntaxGym have enabled systematic syntactic evaluation in English and Spanish, no comparable resource exists for French. To address this gap, we present SyntaxGymFR, a manually curated evaluation suite for evaluating the syntactic abilities of French and multilingual LLMs. SyntaxGymFR consists of manually validated minimal sentence pairs targeting key syntactic phenomena in French. We describe the annotation methodology, the selection of linguistic constructions, and the validation procedures used to ensure the coverage of syntactic phenomena. Furthermore, we report experimental results obtained with several French and multilingual LLMs, analyzing their sensitivity to grammatical contrasts and cross-linguistic transfer effects. Our results provide new insights into the syntactic generalization capabilities of French LLMs and establish SyntaxGymFR as a benchmark for future research on language-specific evaluation of syntactic competence.

MOTS-CLÉS : Évaluation syntaxique, grands modèles de langue, SyntaxGym, annotation de corpus, évaluation comparative.

KEYWORDS: Syntactic Evaluation, Large Language Models, SyntaxGym, Corpus, Benchmarking.

Actes des 33ème Conférence sur le Traitement Automatique des Langues Naturelles.
Volume 2 : articles déjà publiés, pages 48–49.
Cette œuvre est mise à disposition sous licence Attribution 4.0 International.

ARTICLE ACCEPTÉ À : LREC 2026 : The Fifteenth biennial Language Resources and Evaluation Conference, Palma, Mallorca, Spain, May 11-16, 2026.
URL : https://lrec.elra.info/lrec2026-main-178