



**HAL**  
open science

## Un outil d'exploration de corpus d'apprenants pour la didactique des langues secondes

Sarra El Ayari

► **To cite this version:**

Sarra El Ayari. Un outil d'exploration de corpus d'apprenants pour la didactique des langues secondes. Journées d'études " De l'acquisition à la didactique (et vice-versa) ", Nov 2021, Paris, France. hal-03418871

**HAL Id: hal-03418871**

**<https://hal.science/hal-03418871>**

Submitted on 8 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## Contexte

Les recherches en acquisition des langues ont produit (et produisent encore) une multitude de différents types de corpus d'apprenants [5] qui sont transcrits (pour les corpus oraux), annotés et analysés. Beaucoup de ces corpus sont accessibles aujourd'hui dans différentes banques de données (Ortolang, Cocoon, Talkbank par exemple) et prêts à être utilisés et valorisés. Nous proposons une interface qui permet d'utiliser ces corpus d'apprenants en acquisition L2 (oral, écrit) issus de projets de recherche, déjà transcrits et annotés, comme ressources pédagogiques [3] pour l'apprentissage des langues étrangères.

## Corpus

Le Corpus Écrit de Français Langue Étrangère (CEFLE) est un **corpus de textes écrits** en français L2 par des apprenants suédophones [6], [1]. Ce corpus a été élaboré à l'Université de Lund, Suède, durant l'année scolaire 2003/04 et les apprenants du corpus ont été répartis selon leur profil grammatical [2].

## Outil d'exploration d'un corpus d'apprenants

Ce développement se base sur la plateforme en ligne d'annotation et d'exploration de corpus Sarramanka [4].

### Exploration du corpus

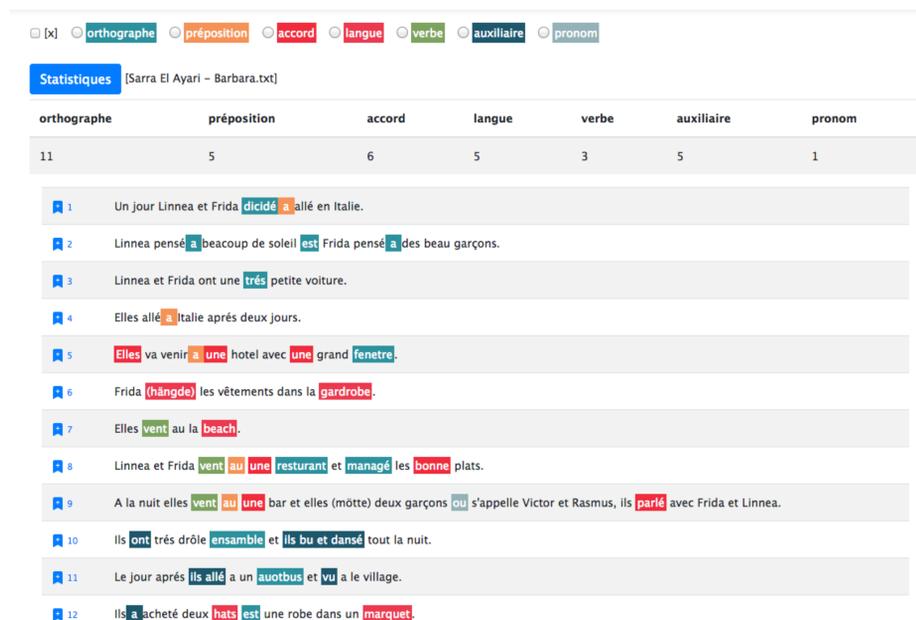


FIGURE 1: Interface de correction des phrases

### Ajout d'une correction

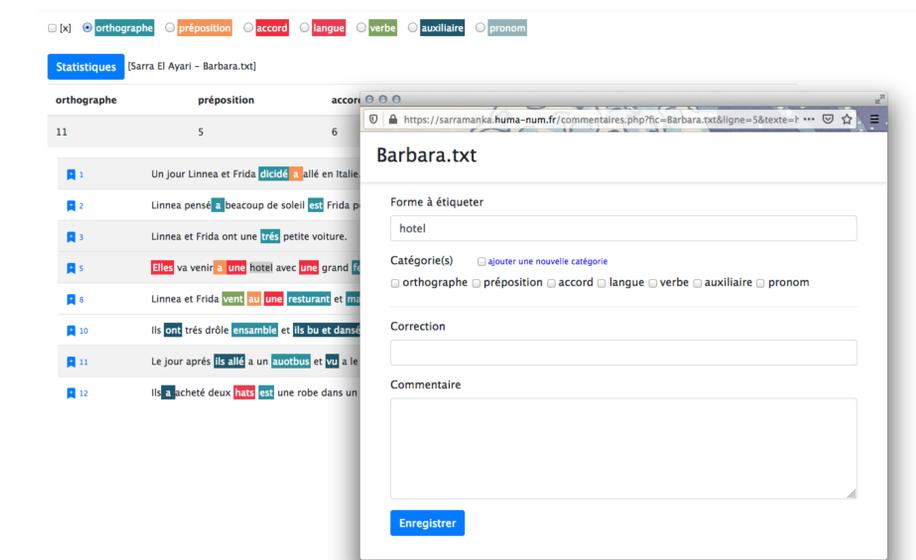


FIGURE 2: Interface d'ajout d'une correction

## Résultats

Apprenant	Orth.	Prép.	Acc.	L1
LBO	5	6	9	2
SEA	7	3	7	2
GIK	8	3	8	2

TABLE 1: Répartition des corrections par apprenants

## Valorisation de corpus et ressources

Les corpus d'apprenants sont en vraie manne en terme de ressources disponibles dans différentes langues. Cette plateforme permet de les rendre accessibles et utilisables à des fins didactiques, et peut être facilement adaptée aux besoins des enseignants en classe, ce qui participe également à leur valorisation dans le cadre des recherches en acquisition des langues qui sont menées (principes FAIR - Findability, Accessibility, Interoperability and Reusability).

## Apports pour la didactique

- travail en classe
- inclusion dans séquence pédagogique
- évaluation
- comparaison des résultats
- export des résultats
- paramétrage par niveau
- suivi de progression (longitudinal)
- accès à la correction

## Apports pour l'acquisition L2

- cheminement de la construction d'un texte écrit
- observation des phénomènes de correction, auto-reformulation
- enregistrement des séquences du processus de l'écriture L2

## Perspectives

- interface personnelle pour l'utilisateur
- accès au texte corrigé
- développement de statistiques et résultats longitudinaux
- mener une recherche/action

## Références

- [1] K. Bardovi-Harlig. Tense and aspect in second language acquisition : Form, meaning and use. *Language Learning*, (50), 2000.
- [2] I. Bartning and S. Schlyter. Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, (14) :281–299, 2004.
- [3] S. De Cock and H. Tyne. Corpus d'apprenants et acquisition des langues. *Recherches en didactique des langues et des cultures*, 1(11), 2014.
- [4] S. El Ayari and M. Watorek. Exploration outillée pour un corpus de productions orales d'apprenants débutants en L2. Colloque Influence translinguistique : où en est-on aujourd'hui? (GIS RéAL2), Toulouse, France 2021.
- [5] S. Granger. Computer learner corpus research : current status and future prospects. In T. Connor, U. & Upton, editor, *Applied corpus linguistics : a multidimensional perspective.*, pages 123–145. Amsterdam/ Atlanta :Rodopi, 2004.
- [6] Ågren M. *À la recherche de la morphologie silencieuse. Sur le développement du pluriel en français L2 écrit.* PhD thesis, Université de Lund, 2008.