



Le Mans
Université



LIUM

Laboratoire d'Informatique
Le Mans Université

LIADILE



UNIVERSITÉ
RENNES 2

Vers des learning analytics linguistiques actionnables pour l'apprentissage de l'anglais en contexte universitaire

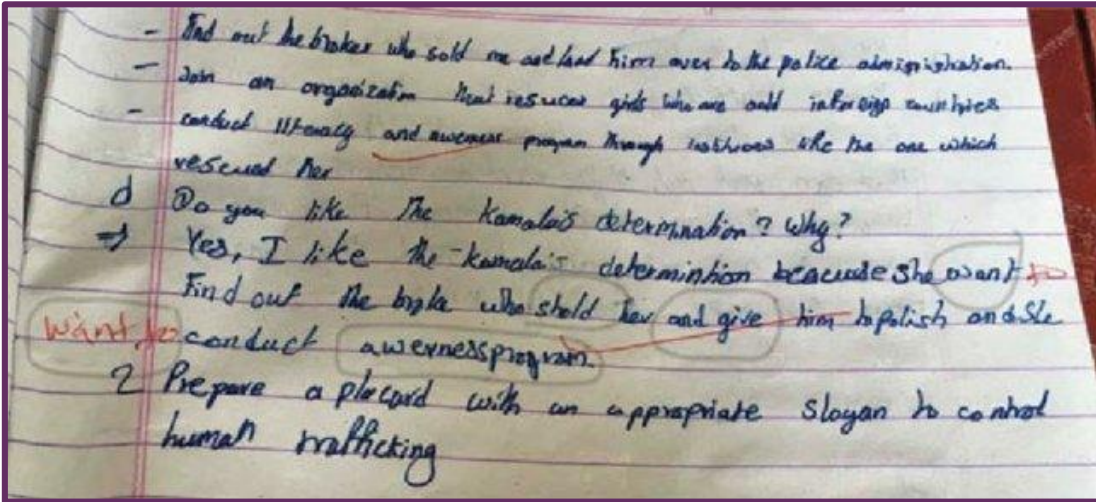
Rémi Venant, Nicolas Baillier, Cyriel Mallart, Andrew Simpkin,
Bernardo Stearns, Jen-Yu Li et Thomas Gaillat



mardi 10 juin 2025

anr[©]

La correction de la langue en pratique



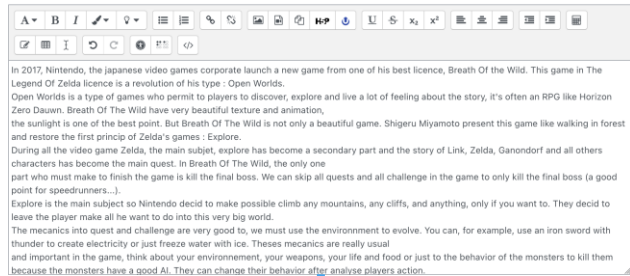
- Correction chronophage.
- Centrée sur les erreurs d'orthographe et de grammaire.
- Sous-estimation d'autres dimensions du langage.

Sah, P. K., & Li, G. (2018). English medium instruction (EMI) as linguistic capital in Nepal: Promises and realities. *International Multilingual Research Journal*, 12(2), 109-123.

A4LL, *Analytics for Language Learning*

- Projet de recherche de 2 ans, financé par l'Agence Nationale de la Recherche.
- Universités de Rennes, Paris, Le Mans, Galway (Irlande) et IRISA/INSA de Rennes.
- Objectif: **créer un système d'analytiques pour l'apprentissage d'une langue étrangère écrite** comme support:
 - À l'objectivation de l'évaluation ;
 - À la prise en compte d'autres dimensions du langage ;
 - À l'évaluation automatique du niveau de maîtrise ;
 - À l'analyse des difficultés de l'apprenant.

Vue d'ensemble de A4LL



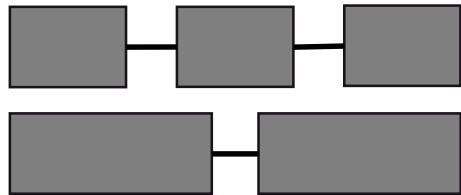
Interactions médiatisées

Classe + Moodle (LTI)



- Prédiction
- Explications
- Texte enrichi

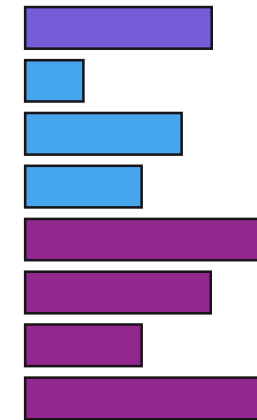
- Flux de texte
- Production finale



Extractions de traits linguistiques



Modèles probabilistes de niveau linguistique



Mesures explicatives des traits linguistiques

Connaissances symboliques

Variables Linguistiques et CECR

- 591 mesures textométriques permettant des comparaison entre textes [1, 2].
 - Différents outils d'extraction utilisés (TAACO, UD_feat, Keylogger).
 - Plusieurs domaines linguistiques ciblés : syntagmatique, paradigmatic, phraséologique, cohésif, comportemental.
- Une variable : non actionnable pour un enseignant.
 - Définition complexe.
 - Sémantique trop éloignée des concepts manipulés et des objectifs pédagogiques.

average lemma frequency, log
transformed - news

Variables Linguistiques et CECR

- CECR : Cadre Européen Commun de Référence en langue [3].
 - Approche par compétence « haut niveau ».
 - Traite l'oral et l'écrit, la compréhension et la production.
 - 6 **niveaux de maîtrise** : A1, A2, B1, B2, C1, C2.
 - Définit des **descripteurs** pour guider l'évaluation.
 - Ex.: *pragmatic competence: thematic development*
 - « Peut rédiger des textes détaillés officiels ou pas sur une gamme étendue de sujets relatifs à son domaine d'intérêt ... »
- Rattachement de chaque variable à un descripteur du CECR.
 - 6 descripteurs CECR retenus : 6 dimensions.
 - *Inégalité de distribution des variables sur les descripteurs.*
 - *Pas de corrélations monotoniques entre variables et niveaux.*

Variables Linguistiques et CECR

Description mesure Bas Niveau	Num. of collocation used	Lda divergence (adjacent paragraphs)	nb backspace sequence longer than 3
Outil	Collocation tool	TAACO	Keylogs
Domaine linguistique	Lexical	semantique	behavioural
Compétence CECR	vocabulary range: diversity	pragmatic competence: thematic development	accuracy

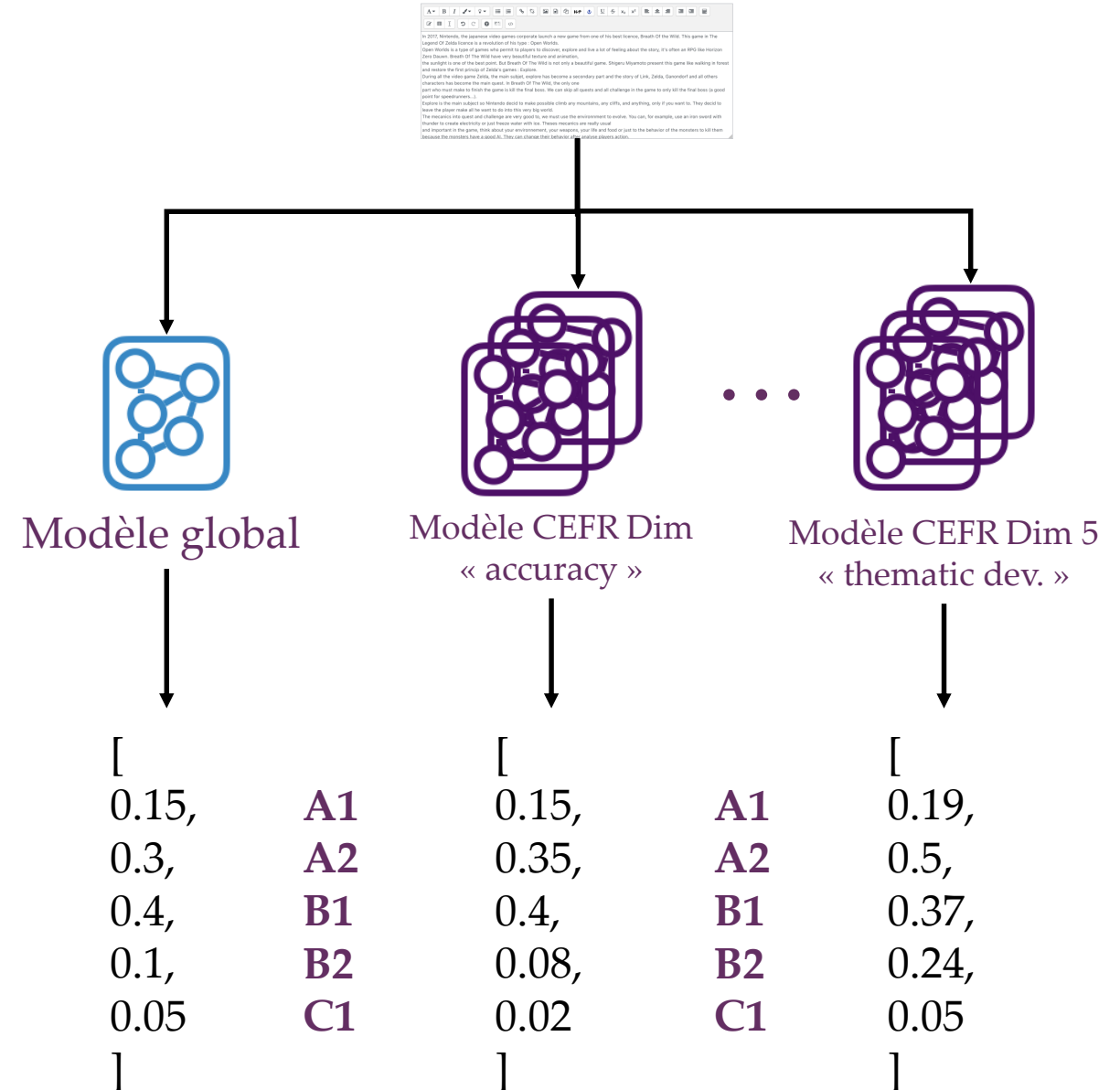
Exemples de variables et de leur métadonnées

Prédire un niveau et l'expliquer : questions de recherche

- Comment sélectionner des mesures pertinentes statistiquement ?
- **Comment lier les mesures linguistiques à des traits langagiers adaptés à l'enseignant ?**
- **Comment créer un Tableau de Bord d'Apprentissage (TBA) pour visualiser et expliquer ces traits et mesures dans un objectif d'actionnabilité ?**

Modèles prédictif

- 2 catégories de classifieurs supervisés.
- Prédiction du niveau CECR.
- Jeu de données labélisés :
 - EFCAMDAT (3/4) : 5238.
 - CELVA [4] (1/4) : 1733.
- Modèle 1 : prédiction du niveau sur l'ensemble des variables.
 - Elastic Net + Lasso.
 - Bonnes performances (> 63% précision).
- Modèles 2 : prédiction par dimension CECR.
 - 5 modèles ordinaux Random Forest.
 - Modèles moins performants mais plus explicables.



Méthode d'élaboration du tableau de bord

- Nécessité de prendre en compte l'enseignant dans la conception du TBA.
- Exploitation et adaptation de la méthode **PADDLE** [5].
 - Co-concevoir un TBA en identifiant son objectif, son contexte d'utilisation, le choix des données et des visualisations, et élaborer plusieurs écrans de ce dernier.
- Spécificités du contexte : les variables et modèles préexistent au TBA :
 - **appropriation** des données et non pas conception.
 - **Dérivation des objectifs** à partir de cette base.
- **10 enseignantes** d'Anglais à l'Université Rennes 2.

Session 1

Présentation du projet et
du modèle de variables
Session PADDLE

3h

Evaluation de la compréhension.
Cas d'usage ; Visualisation.
UI/UX globale.

Dév.

Maquette
fonctionnelle

Session 2

Evaluation de la
maquette

3h

Evaluation par cas d'usage.
Evaluation de l'interprétation.
Evaluation de l'actionnabilité.

Dév.

TBA
opérationnel

Exploitation
en contexte
réel

Session 3

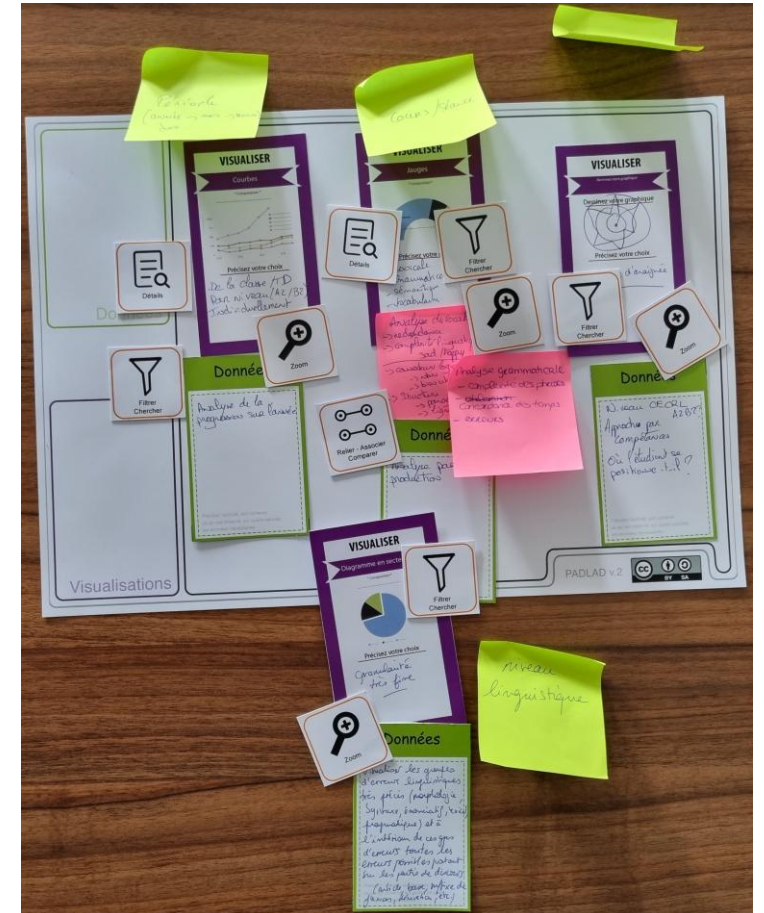
Evaluation du
TBA

3h

Evaluation de la compréhension.
Cas d'usage ; Visualisation.
UI/UX globale.

Conclusion des deux premières sessions

- **Changement de paradigme** d'évaluation **difficile**.
 - Pas focalisé sur les fautes.
 - Dérive au fur et à mesure de la première séance.
- **Capture délicate** du lien entre cas d'usage et **parcours utilisateur** du TBA.
 - Non linéarité du parcours.
- Objectifs retenus
 - Identifier le **niveau de l'étudiant**.
 - Identifier **des éléments de feedback positifs**.
 - Permettre la montée en compétence selon **les faiblesses identifiées** en langue (fluidité, structure, aisance).
 - Permettre **une remédiation auto. par la génération d'exercices**.
 - Permettre la création d'ateliers sur mesure avec des ressources et des activités personnalisées.
- Choix des visualisation.
 - **Appropriation vs. Avis expert**.
 - Principales retenues : histogramme, radar.



TBA développé : vue d'ensemble

Sports - Exercice 1 Prêt

Activity Dashboard

Overview

Unselect all

Henry Descartes (B1)

Text

Bob Descartes (A2)

Text

Henry Freville (B1)

Text

Julie Fuye (A2)

Text

Bernardo Jolivet (B1)

Text

Henry Luengo (A2)

Text

Bernardo Bellanger (A2)

Text

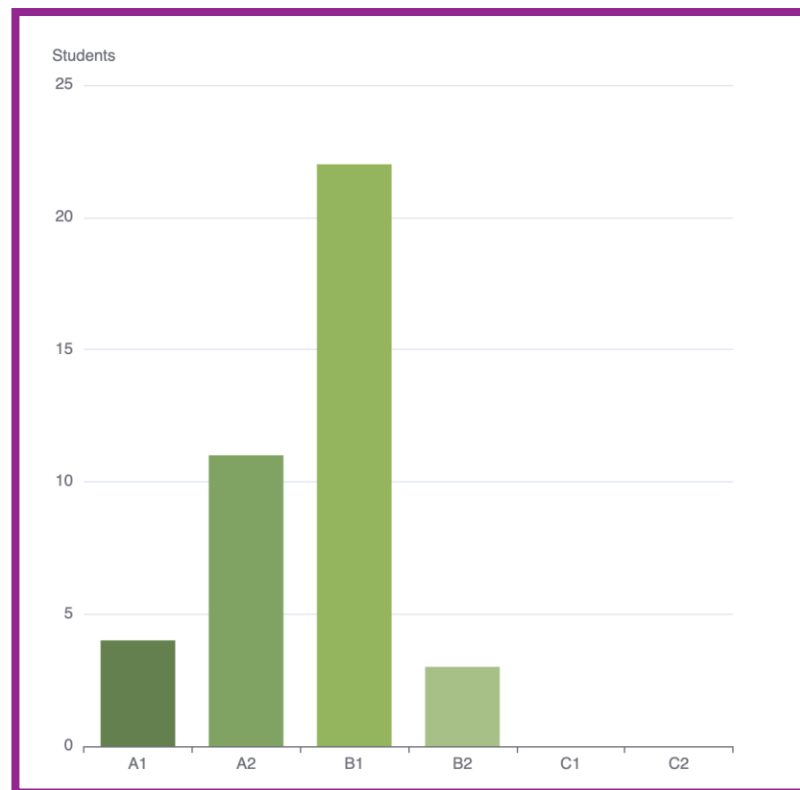
Jean Luengo (B1)

Text

Bob Velpau (B1)

Text

Distribution of students by overall CEFR level



CEFR Predictions distribution

On this visualization, you can see the distribution of the CEFR levels of each students' writing computed by a global predictive model. The students' writings were compared with those of a reference cohort based on many linguistic measures.

For each writing, the model computes the probabilities of being assigned to each CEFR level. The level with the highest probability is selected. To assess the confidence of the predictions, select one or several students, and check the probabilities of scores. This indicates how confident you can be about the score.

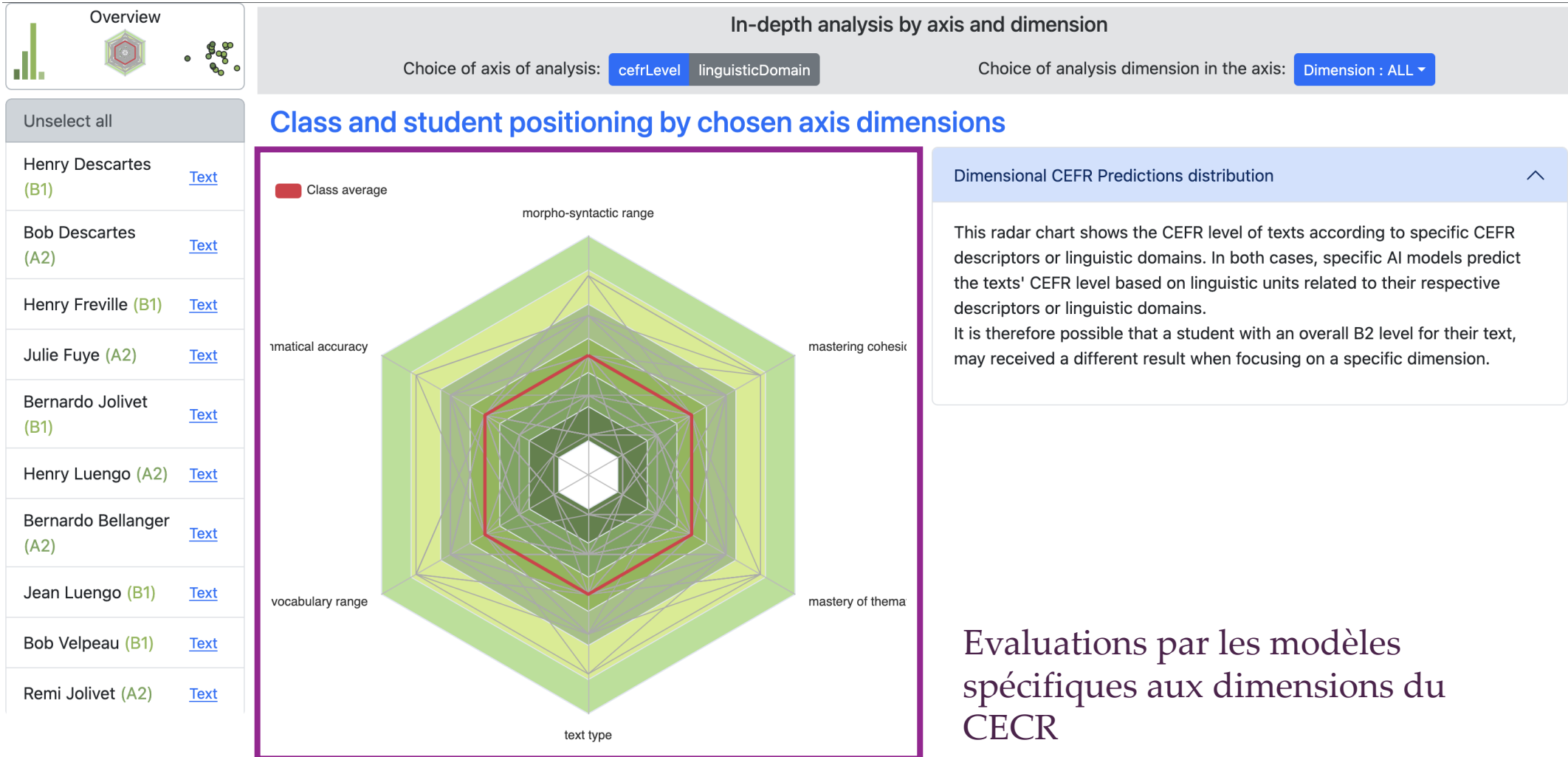
If you select a single student you will also obtain some explanations about the most salient linguistic information that pushed the model's decision.

Panneau d'explication

Navigation et sélection

Distribution selon le modèle global

TBA développé : vue d'ensemble



Evaluations par les modèles spécifiques aux dimensions du CECR

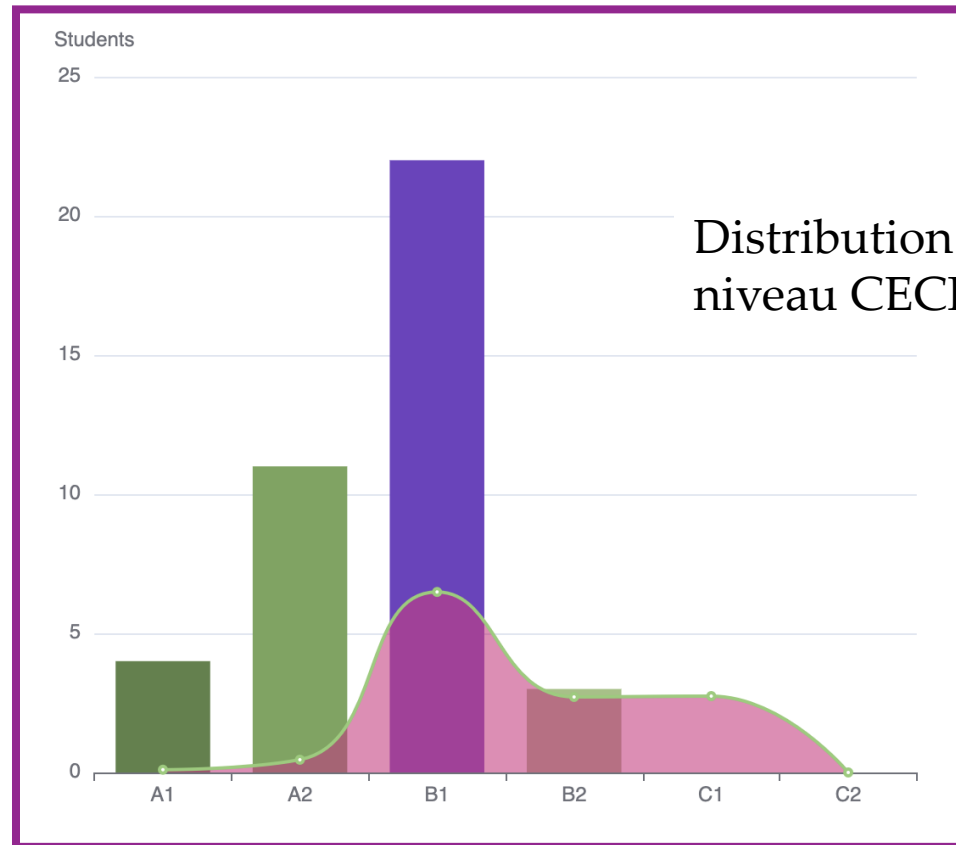
TBA : exploration par étudiant(s)

Overview

Unselect all

Henry Descartes (B1)	Text
Bob Descartes (A2)	Text
Henry Freville (B1)	Text
Julie Fuye (A2)	Text
Bernardo Jolivet (B1)	Text
Henry Luengo (A2)	Text
Bernardo Bellanger (A2)	Text
Jean Luengo (B1)	Text

Distribution of students by overall CEFR level



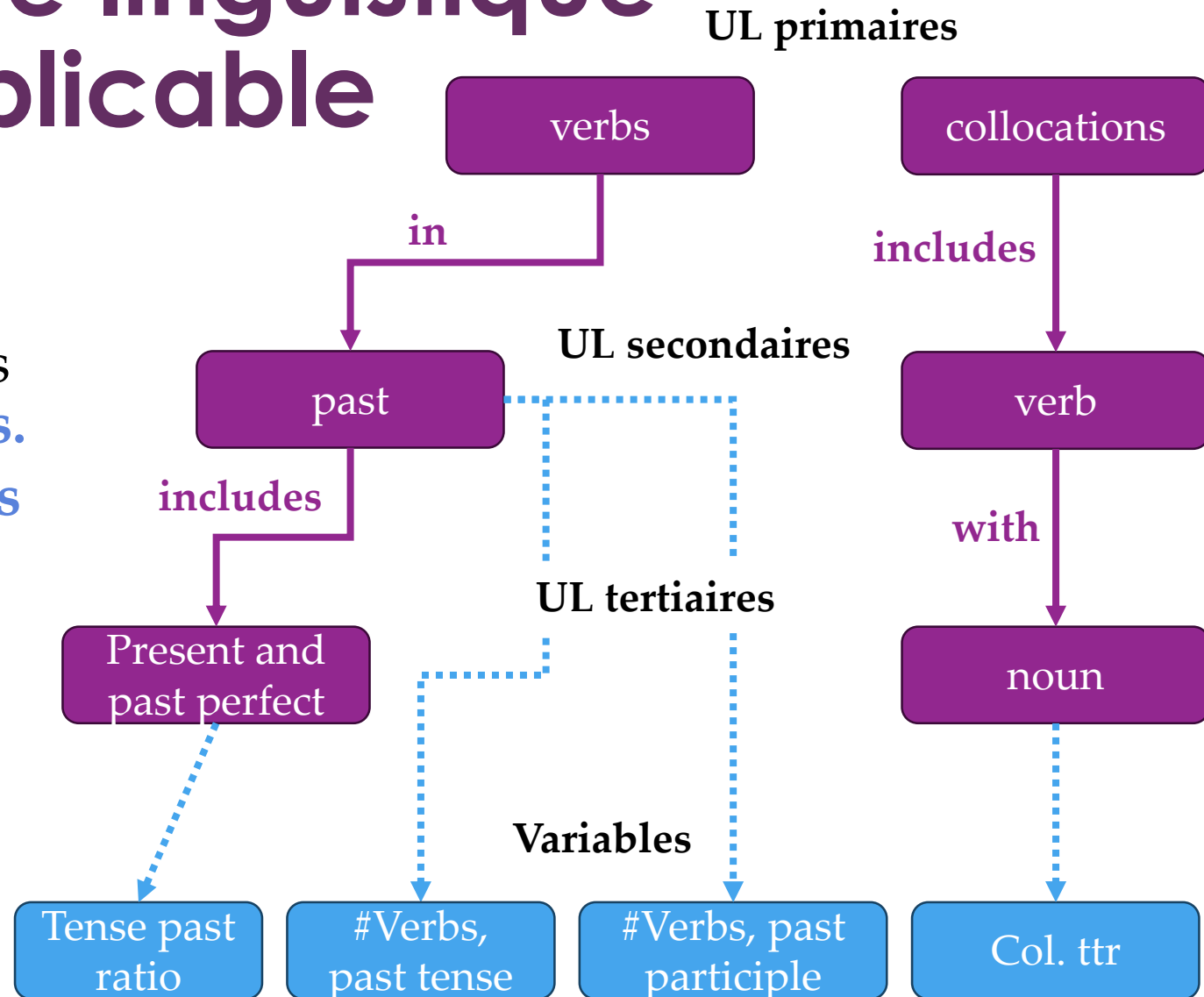
Distribution des probabilités de niveau CECR pour l'étudiant

Expliquer une prédiction

- L'affichage des **distributions de probabilité** de niveau du modèle global et des prédictions par dimensions de CECR :
 - permet une **évaluation fine** du niveau prédit,
 - mais reste **trop abstrait** pour permettre à l'enseignant de prendre une décision de remédiation.
- Travailler à **l'échelle des variables** pour expliquer les décisions des modèles :
 - trop « bas niveau » : **variables trop nombreuses** avec des **sémantiques inadaptées** à l'enseignant.
- Travailler à **l'échelle des dimensions du CECR**
 - Utile pour **l'exploration et l'affinage des prédictions**
 - Reste **trop abstrait** pour prendre une décision de remédiation.

Explications : unité linguistique comme pivot explicable

- Proposition [6]
 - Une ontologie hiérarchique d'unités linguistiques reliées par des connecteurs logiques.
 - notions conceptuelles propres au métier d'enseignant de langues étrangères.



Explications : unité linguistique comme pivot explicable

UL primaires

verbs

collocations

in

includes

UL secondaires

past

verb

with

noun

Quelles unités linguistiques pour expliquer une prédiction faite sur un texte d'un apprenant ?

Variables

Tense past ratio

#Verbs, past tense

#Verbs, past participle

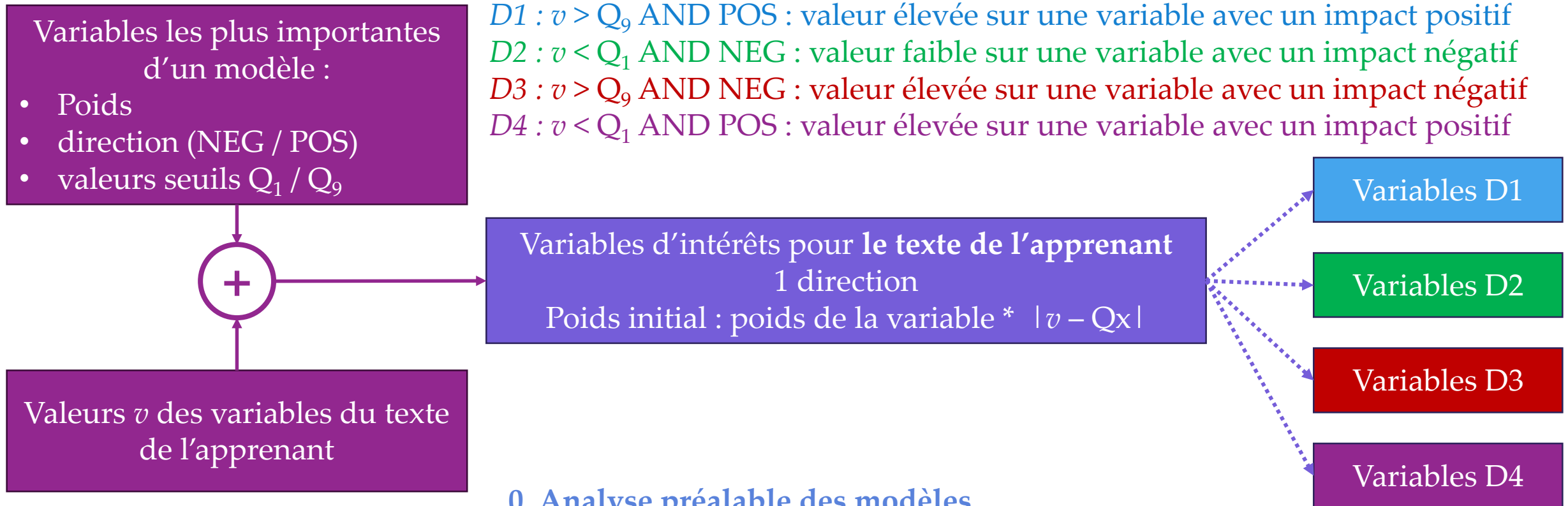
Col. ttr

- Problématiques :
 - Une variable n'est pas compréhensible pour l'enseignant
 - Les dimensions CECR sont utilisées pour l'exploration des prédictions, mais trop étroites pour fournir
- Pro
 - d'unités linguistiques reliées par des connecteurs logiques
 - notions conceptuelles propres au métier d'enseignant de langues étrangères.

Moteur générique d'explication

Directions

- $D1 : v > Q_9$ AND POS : valeur élevée sur une variable avec un impact positif
 $D2 : v < Q_1$ AND NEG : valeur faible sur une variable avec un impact négatif
 $D3 : v > Q_9$ AND NEG : valeur élevée sur une variable avec un impact négatif
 $D4 : v < Q_1$ AND POS : valeur élevée sur une variable avec un impact positif



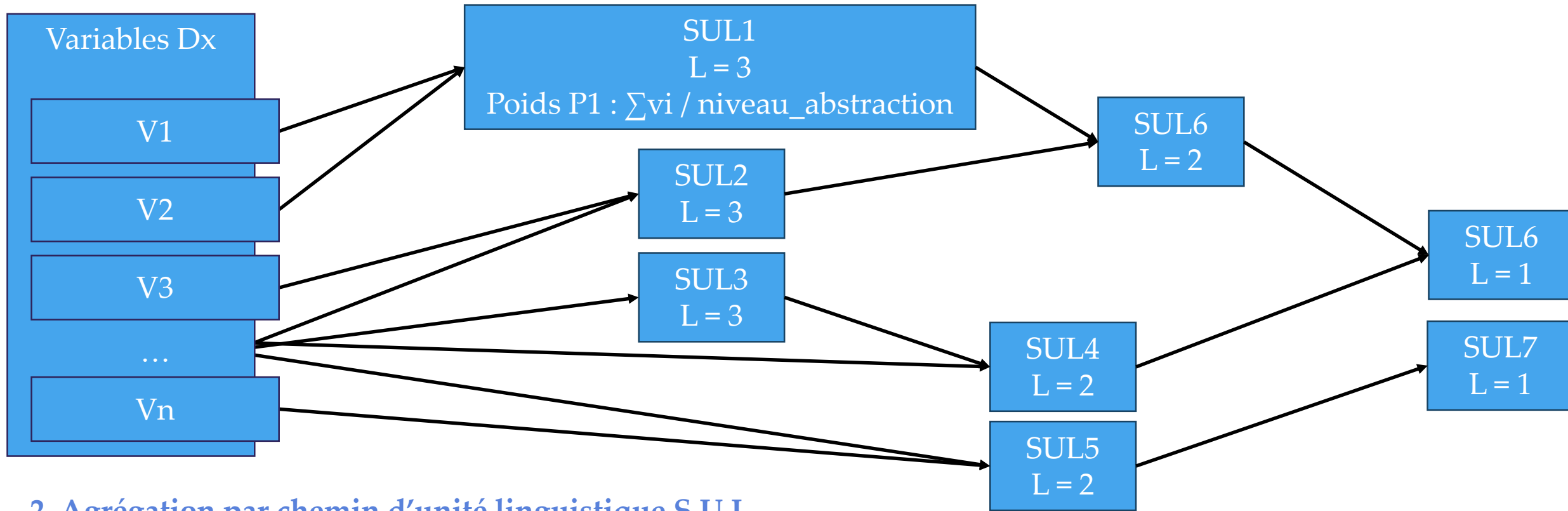
0. Analyse préalable des modèles

Ex. pour les RF : Permutation Importance / Partial Dependence

Analyse préalable des distributions statistiques des valeurs (déciles 1 et 9)

1. Sélections des variables d'intérêt pour l'étudiant

Moteur générique d'explication

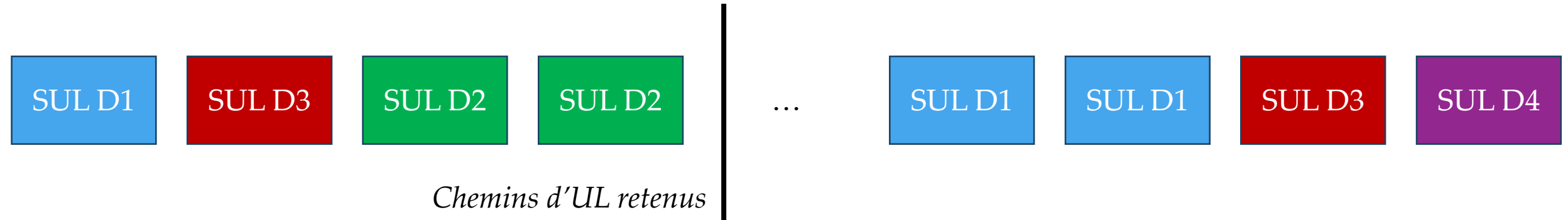


2. Agrégation par chemin d'unité linguistique S.U.L.

Equilibrage du poids d'un SUL

- **Augmente** avec le nombre de variable regroupés (sommées des poids)
- **Diminue** avec le niveau d'abstraction du chemin (plus le chemin est court, plus le trait linguistique est abstrait)

Moteur générique d'explication



3. Sélection des chemins d'UL les plus pertinents pour expliquer la décision du modèle

- Les **SUL des 4 direction sont triés** par poids décroissant
- Les **SUL « contradictoires » sont éliminés** (même chemin pour des directions différentes)
- Les **k** premiers sont retenus, k paramètre arbitraire

Généralisation

- Algorithme **généralisé pour s'appliquer sur un ensemble d'apprenant**
 - Sélection des SUL commune à tous les apprenants du groupe ou bien à un % supérieur à un seuil donné

Limitations

- *Comment définir k ?*
- *Risques de contradictions résiduels ?*

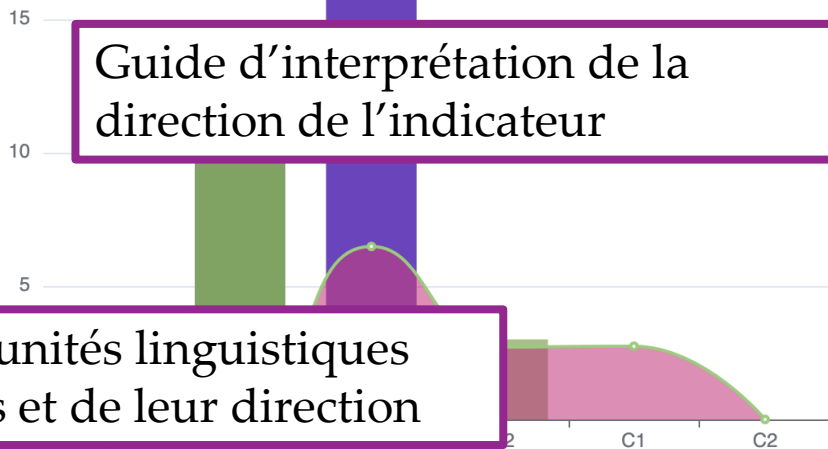
Restitution des explications

Indicateur : séquence d'unité linguistique sélectionnée

Accès au détails des mesures

Guide d'interprétation de la direction de l'indicateur

Résumé des unités linguistiques sélectionnées et de leur direction



Indicator	words re...	content ...	nouns/pr...	content ...	nouns re...
Value	↓	↑	↑	↑	↓
Impact	↓	↓	↑	↑	↓

Indicator 2: content words varying within texts

The indicator is a grouping of 1 measures which combine **content words varying within texts**.

[Show the list of measures](#)

Interpretation:

The value of this indicator is high in the student's writing. The indicator is associated with worse proficiency.

⬇ Try to see why the use of content words varying within texts drives the student to lower proficiency.

Generate pedagogical activity

Indicator 3: nouns/pronouns repeating in consecutive sentences

The indicator is a grouping of 2 measures which combine **nouns/pronouns repeating in consecutive sentences**.

[Show the list of measures](#)

Génération d'activité à partir des chemins d'unité linguistiques

Generate LLM Based pedagogical activity



According to this indicator and the student(s) selected, the available activities are:

- ☐ Cloze based on generated content
- ☐ Translation from generated content
- ☒ Cloze based on student's writing
- ☐ Rephrasing based on student's writing
- ☐ Matching based on original content
- ☐ Chat based on original content

Generate Activity

Choix de l'activité

Cloze based on student's writing

A cloze exercise focusing on the current indicator, built from the student's text.

You are a teacher of English as a foreign language. You want to take the student's text <I'll give to you my opinion about VAR VAR had been introduced in football almost 6-8 years ago. At t Today, after a lot of games played under VAR control, we can s In my opinion, even with VAR using during football games, it s For all those facts, I think that VAR isn't verry well used in For example. last week end. OM defeated FC Nantes in Paris cha

Click on the copy button in the top right corner of the prompt above or double click anywhere on the text to select it. Click on the copy button in the top right corner of the prompt or double click anywhere on the prompt to select it.

You can then paste it to an generative AI specialized in english training. As an example you can use <https://chatgpt.com/g/g-rJyYgZ0cr-correct-my-english>

Prompt adapté s'appuyant sur l'indicateur (séquence d'unité linguistique) choisie

Proposition de modèle approprié

Close

Evaluation du tableau de bord (Session 3)

- 10 enseignantes, 3h
- Objectifs d'évaluation
 - Pertinence des recommandations fondés sur les unités linguistiques.
 - Actionnabilité des recommandations.
 - Expérience utilisateur.
- Activité
 - Réalisation de tâches de diagnostics sur des apprenants précis issus de contexte réel.
 - User Experience Questionnaire.
 - Table ronde : retour d'expérience et commentaire.
- Représentation graphique et navigation **claires et intuitive**
- Intégration avec Moodle très appréciée
- Assistance à la **génération d'activité considérée très utile**
- **Charge cognitive importante** sur les explications : trop d'informations à l'écran, scrolling nécessaire
- Terminologie des **UL pas toujours intuitive**
- Résultat du **UEQ bons** :
 - Attraction, Efficacité, Contrôlabilité, Stimulation et originalité bons (> 0,9)
 - Compréhensibilité neutre (0,17)

Conclusion et perspectives

- Comment lier les mesures linguistiques à des traits langagiers adaptés à l'enseignant ?
 - **Ontologie hiérarchique d'unités linguistiques adaptés au vocabulaire enseignant**
- Comment créer un Tableau de Bord d'Apprentissage pour visualiser et expliquer ces traits et mesures dans un objectif d'actionnabilité ?
 - Méthode **PADDLE centrée utilisateur adaptée et étendue**
 - **Moteur générique d'explication** fondée sur l'analyse des variables importantes propres à l'apprenant et sur **l'extraction d'unités linguistiques pertinentes** à l'explication
 - Proposition de **génération automatique d'activités** fondées sur les unités linguistiques extraites
 - Evaluations multiples sur le TBA dans son ensemble et sur les explications elles-mêmes

Perspectives

- **Court et moyen terme**

- **Validation de la taxonomie** : mesure de l'accord inter-juge entre **linguistes** (en cours).
- **Intégration des LLM** comme source de **prédiction du niveau** (en cours).
- **Visualisation** des indicateurs au **niveau du texte**, en particulier la **dynamique de production** du texte.
- Amélioration des modèles par **l'enrichissement du Corpus Celva**.

- **Long terme**

- Adaptation des modèles à **d'autres langues** (collecte en cours pour l'espagnol et l'allemand).
- Prise en compte des **production orales**.
- Etude **d'impact sur l'apprentissage et l'enseignement**.

Vers des learning analytics linguistiques actionnables pour l'apprentissage de l'anglais en contexte universitaire

Merci pour votre attention !



LIUM
Laboratoire d'Informatique
Le Mans Université



LIADILE



Références

1. K. Kyle, S. Crossley, & C. Berger., « The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0 », Behavior Research Methods, pp. 1030–1046, 2018.
2. T. Gaillat et al., « Predicting CEFR levels in learners of English: The use of microsystem criterial features in a machine learning approach », ReCALL, vol. 34, no. 2, pp. 130–146, 2022.
3. Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. Common European framework of reference for languages: Learning, teaching, assessment. Cambridge University Press, 2001.
4. Gaillat, T., Venant, R., Mallart, C., Arnold, T., & Fougère, A. CELVA. Sp corpus User Interface (Doctoral dissertation, Université Rennes 2; Université de Rennes), 2024.
5. Jean-Marie Gilliot et al. « Conception participative de tableaux de bord d'apprentissage », IHM'18: 30e Conférence Francophone sur l'Interaction Homme-Machine, pp-19, 2018.
6. Mallart, C., Simpkin, A., Venant, R., Ballier, N., Stearns, B., Yu-Li, J., & Gaillat, T., « Vers une grammaire probabiliste de microsystèmes fonctionnels en L2 », RéAL2: Grammaire (s) et acquisition des L2: Approches, trajectoires, interfaces, 2023.