

Les bases d'un dialogue entre les didacticien·ne·s et les statisticien·ne·s

Vincent Brault

sur des travaux en collaboration avec Sylvain Coulangue, Margaux Leroy,
Frédérique Letué, Marie-Pierre Jouannaud, Marie-Jo Martinez et
Anne-Cécile Perret

Jeudi 15 octobre



LABORATOIRE
JEAN KUNTZMANN
MATHÉMATIQUES APPLIQUÉES - INFORMATIQUE

CoPoLang

UGA
Université
Grenoble Alpes

Plan

- 1 Introduction
- 2 Collaborations : exemple du test SELF
- 3 Approche par classification
- 4 Programme de la journée

Qu'est-ce qu'un test ?

Plusieurs objectifs :

- Certifier des compétences et un niveau acquis, par le biais d'un diplôme (valable à vie).
Ex : *DELFI/DALFI, DFI, DUEFI, First Certificate, CLES...*
- Certifier des compétences et un niveau acquis, par le biais d'un test de niveau (à renouveler).
Ex : *TCFI, TEFI, TOEIC...*
- Positionner un étudiant dans un groupe adapté pour son apprentissage, par le biais d'un test.
Ex : *SELF...*
- Diagnostiquer les forces et les faiblesses linguistiques, dans le but de l'aider dans son apprentissage.

Qu'est-ce qu'un test ?

Plusieurs objectifs :

- Certifier des compétences et un niveau acquis, par le biais d'un diplôme (valable à vie).
Ex : *DELFI/DALFI, DFP, DUEF, First Certificate, CLES...*
- Certifier des compétences et un niveau acquis, par le biais d'un test de niveau (à renouveler).
Ex : *TCF, TEF, TOEIC...*
- Positionner un étudiant dans un groupe adapté pour son apprentissage, par le biais d'un test.
Ex : *SELF...*
- Diagnostiquer les forces et les faiblesses linguistiques, dans le but de l'aider dans son apprentissage.

Qu'est-ce qu'un test ?

Plusieurs objectifs :

- Certifier des compétences et un niveau acquis, par le biais d'un diplôme (valable à vie).
Ex : *DELFDALF, DFP, DUEF, First Certificate, CLES...*
- Certifier des compétences et un niveau acquis, par le biais d'un test de niveau (à renouveler).
Ex : *TCF, TEF, TOEIC...*
- Positionner un étudiant dans un groupe adapté pour son apprentissage, par le biais d'un test.
Ex : *SELF...*
- Diagnostiquer les forces et les faiblesses linguistiques, dans le but de l'aider dans son apprentissage.

Qu'est-ce qu'un test ?

Plusieurs objectifs :

- Certifier des compétences et un niveau acquis, par le biais d'un diplôme (valable à vie).
Ex : *DELFI/DALFI, DFI, DUEFI, First Certificate, CLES...*
- Certifier des compétences et un niveau acquis, par le biais d'un test de niveau (à renouveler).
Ex : *TCFI, TEFI, TOEIC...*
- Positionner un étudiant dans un groupe adapté pour son apprentissage, par le biais d'un test.
Ex : *SELF...*
- Diagnostiquer les forces et les faiblesses linguistiques, dans le but de l'aider dans son apprentissage.

Exemple du test SELF



Tourist brochure



One of the more surprising aspects of the drive for first-time visitors is that there are virtually no information or directional signs along the way; this is all natural beauty out here, with no signs to spoil the view. So, to fully enjoy yourself and discover all there is to see and do along the drive, it's suggested that you pick up the CD tour of the Hana coast – it's well worth it. They're located only at the Shell Service Station on Route 380 in Kahului.

1/1



According to the brochure, why are there no information signs on the road?

- Because the local people can't afford them.
- Because information signs are ugly.
- Because they want tourists to buy the CD.

10. *Journal of the American Medical Association*, 2000; 283: 2689-2694.



10. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 10, 1175-1182.





Recette de cuisine sur votre blog



Soupe à l'oignon

- les oignons en rondelles
- faire cuire les oignons dans une casserole avec du beurre et de l'huile puis de la farine
- verser le vin blanc dans la casserole
- couvrir la casserole pendant 1h30 et mélanger de temps en temps en tournant avec une cuillère

Bon appétit ! 🍴

A A E E E I I O U U O a a e e e i i o o o

Exemple du test SELF



⏮ ⏪ ⏩ ⏭

Mail ad un'amica

Oggetto: Saluti da Brindisi

Cara Sandra,

Ti scrivo dalla bellissima Brindisi, una città che si trova nella costa est della Puglia. Sono arrivata da una settimana e ogni giorno cose meravigliose.

L'altra sera passeggiavo sul lungo mare e ho trovato una gelateria che fa il gelato del mondo!

Per non parlare dell' pizza che fanno proprio di fronte all'hotel! Domani vado a vedere il primo appartamento e una ragazza mi ha detto che, di norma, le spese per i servizi sono minime.

Ti abbraccio, Carla.

A E E I I O O U a e e i i o o

Exemple du test SELF



- 6 langues (Anglais, Français Langue Etrangère (FLE), Japonais, Espagnol, Italien et Mandarin),
- en moyenne 60 minutes,
- test semi-adaptatif,
- compréhensions de l'oral, de l'écrit et expression écrite courte.

Exemple du test SELF



- 6 langues (Anglais, Français Langue Etrangère (FLE), Japonais, Espagnol, Italien et Mandarin),
- en moyenne 60 minutes,
- test semi-adaptatif,
- compréhensions de l'oral, de l'écrit et expression écrite courte.

Exemple du test SELF



- 6 langues (Anglais, Français Langue Etrangère (FLE), Japonais, Espagnol, Italien et Mandarin),
- en moyenne 60 minutes,
- test semi-adaptatif,
- compréhensions de l'oral, de l'écrit et expression écrite courte.

Exemple du test SELF



- 6 langues (Anglais, Français Langue Etrangère (FLE), Japonais, Espagnol, Italien et Mandarin),
- en moyenne 60 minutes,
- test semi-adaptatif,
- compréhensions de l'oral, de l'écrit et expression écrite courte.

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡

Comment proposer un test *valide* et *fiable* ?²

Un cahier des charges complexe :

- Un temps adapté : si trop long, risque de lassitude chez les candidat·e·s.
- Des questions discriminantes : de telle sorte que des personnes d'un même niveau répondent de façon similaire.
- Le plus automatique possible sans intervention humaine.
- Être le plus authentique possible.
- évaluer la langue selon les critères du cadre européen commun de référence
- Faire attention aux risques de fuite des questions et des réponses¹

¹ 14h30 - **Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?

² 11h00 - **Sébastien Georges** : Le tout est plus que la somme des parties ◀ ≡ ▶ ≡ ↺ 🔍 ↻

Plan

- 1 Introduction
- 2 Collaborations : exemple du test SELF
- 3 Approche par classification
- 4 Programme de la journée

Test à étapes

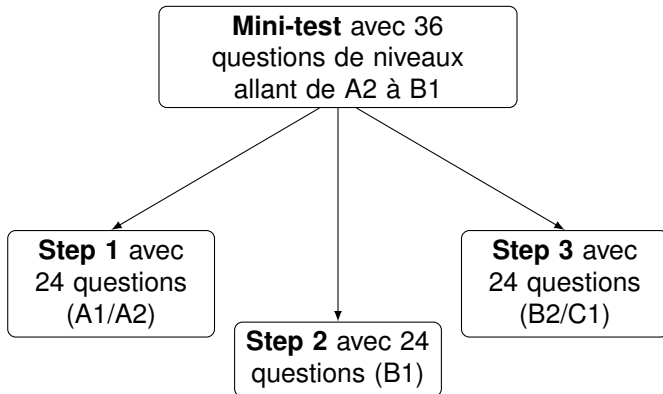
SELF Anglais



Un niveau est donné à la fin en fonction d'un score et suivant un nombre de bonnes réponses dans des groupes d'items.

Test à étapes

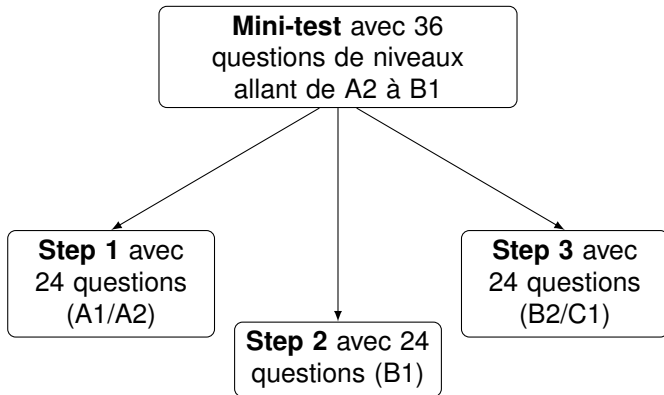
SELF Anglais



Un niveau est donné à la fin en fonction d'un score et suivant un nombre de bonnes réponses dans des groupes d'items.

Test à étapes

SELF Anglais



Un niveau est donné à la fin en fonction d'un score et suivant un nombre de bonnes réponses dans des groupes d'items.

Test à étapes

Résultats pour le test **SELF Japonais**

[Exporter au format PDF](#)

 Groupe cible conseillé

En route vers A2.1

 Compréhension de l'oral

A1.2

 Compréhension de l'écrit

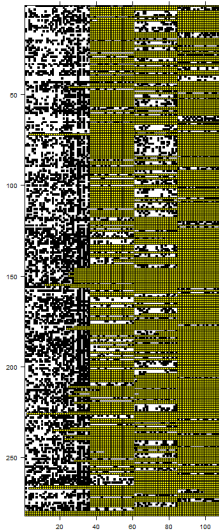
A1.2

 Expression écrite courte

A1.2

Tableau de réponses

Données brutes



Juste



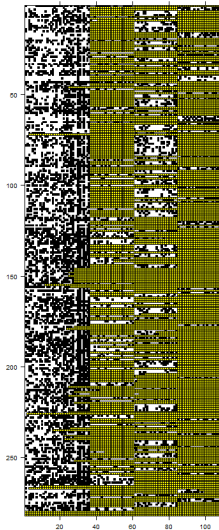
Faux



Non-réponse

Tableau de réponses

Données brutes



Juste



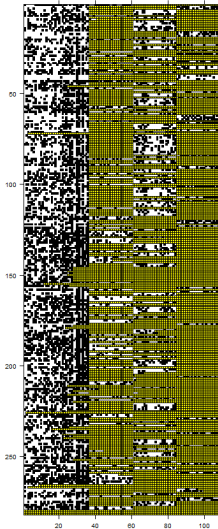
Faux



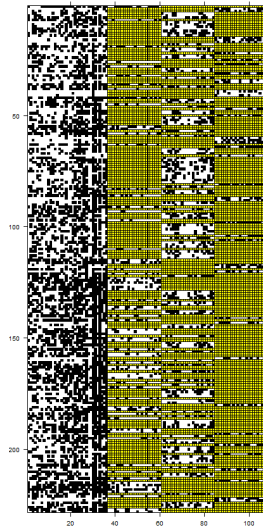
Non-réponse

Tableau de réponses

Données brutes



Elimination des *non-finis*



Juste



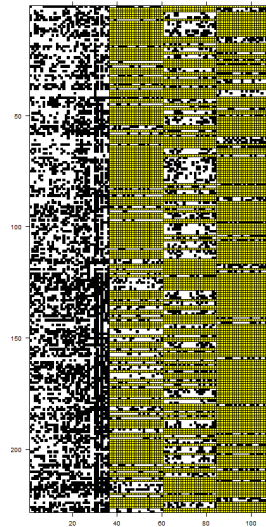
Faux



Non-réponse

Tableau de réponses

Elimination des *non-finis*



Juste



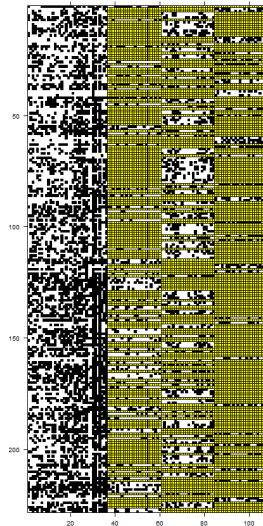
Faux



Non-réponse

Tableau de réponses

Elimination des *non-finis*



Juste



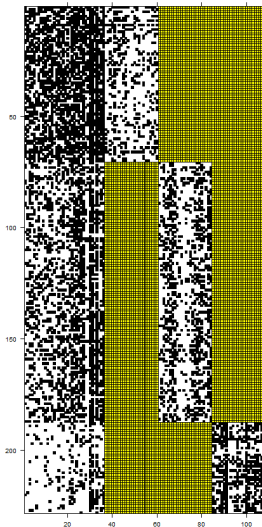
Faux



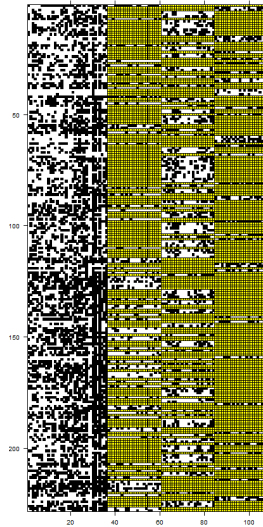
Non-réponse

Tableau de réponses

Données réorganisées



Elimination des *non-finis*



Juste



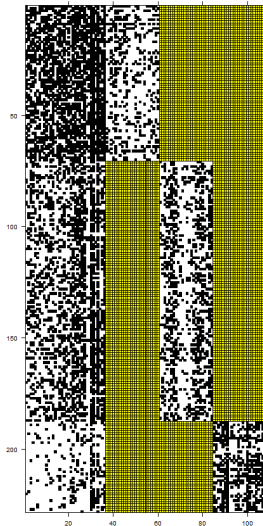
Faux



Non-réponse

Tableau de réponses

Données réorganisées



Juste



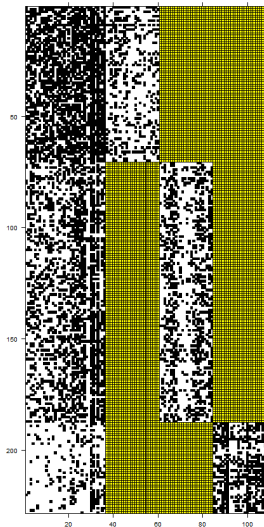
Faux



Non-réponse

Tableau de réponses

Données réorganisées



Juste



Faux



Non-réponse

Questionnement

Question typique à se poser lors de la création d'un test :

- Résultats en accord avec les groupes qui auraient été faits par d'autres méthodes ?
- Calibration des questions pour une efficacité optimale.
- Influence du test initial (ou commun) dans les résultats finaux.

Questions dans le projet CoPoLanguages :

- Peut-on utiliser les résultats des étudiants aux autres *step* pour mieux estimer les groupes ?
- Influences de co-variables comme les langues maternelles ?

Questionnement

Question typique à se poser lors de la création d'un test :

- Résultats en accord avec les groupes qui auraient été faits par d'autres méthodes ?
- Calibration des questions pour une efficacité optimale.
- Influence du test initial (ou commun) dans les résultats finaux.

Questions dans le projet CoPoLanguages :

- Peut-on utiliser les résultats des étudiants aux autres *step* pour mieux estimer les groupes ?
- Influences de co-variables comme les langues maternelles ?

Questionnement

Question typique à se poser lors de la création d'un test :

- Résultats en accord avec les groupes qui auraient été faits par d'autres méthodes ?
- Calibration des questions pour une efficacité optimale.
- Influence du test initial (ou commun) dans les résultats finaux.

Questions dans le projet CoPoLanguages :

- Peut-on utiliser les résultats des étudiants aux autres *step* pour mieux estimer les groupes ?
- Influences de co-variables comme les langues maternelles ?

Questionnement

Question typique à se poser lors de la création d'un test :

- Résultats en accord avec les groupes qui auraient été faits par d'autres méthodes ?
- Calibration des questions pour une efficacité optimale.
- Influence du test initial (ou commun) dans les résultats finaux.

Questions dans le projet CoPoLanguages :

- Peut-on utiliser les résultats des étudiants aux autres *step* pour mieux estimer les groupes ?
- Influences de co-variables comme les langues maternelles ?

Questionnement

Question typique à se poser lors de la création d'un test :

- Résultats en accord avec les groupes qui auraient été faits par d'autres méthodes ?
- Calibration des questions pour une efficacité optimale.
- Influence du test initial (ou commun) dans les résultats finaux.

Questions dans le projet CoPoLanguages :

- Peut-on utiliser les résultats des étudiants aux autres *step* pour mieux estimer les groupes ?
- Influences de co-variables comme les langues maternelles ?

Plan

- 1 Introduction
- 2 Collaborations : exemple du test SELF
- 3 Approche par classification
- 4 Programme de la journée

Homogénéité des groupes

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	
Etudiant 1								→ 4/7
Etudiant 2								→ 4/7

→ Mêmes scores mais des profils différents

Homogénéité des groupes

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	
Etudiant 1								→ 4/7
Etudiant 2								→ 4/7

→ Mêmes scores mais des profils différents

Homogénéité des groupes

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	
Etudiant 1								→ 4/7
Etudiant 2								→ 4/7

→ Mêmes scores mais des profils différents

Homogénéité des groupes

	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	
Etudiant 1	■			■		■		→ 4/7
Etudiant 2		■	■				■	→ 4/7

→ Mêmes scores mais des profils différents

Modèle des blocs latents³

Tableau avant réorganisation

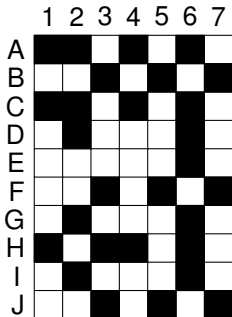


Tableau après réorganisation

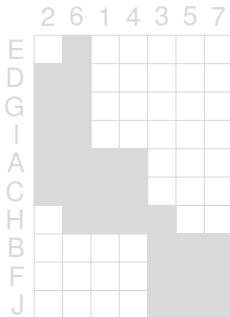
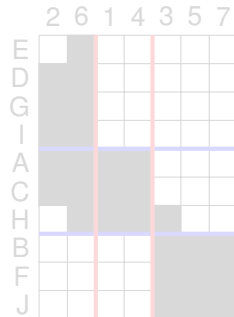


Tableau avec co-classification



³Govaert and Nadif [2003]

Modèle des blocs latents³

Tableau avant réorganisation

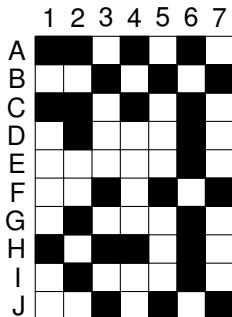


Tableau après réorganisation

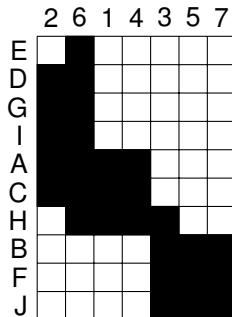
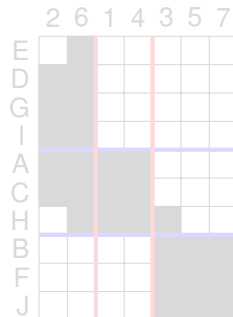
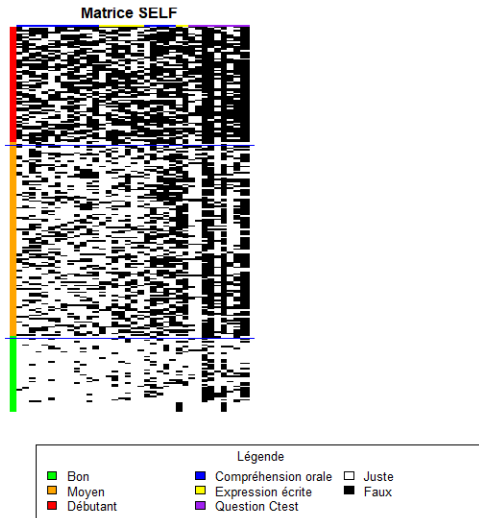


Tableau avec co-classification

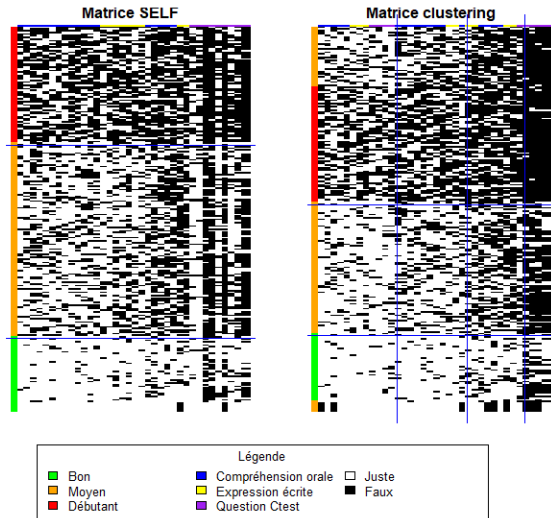
³Govaert and Nadif [2003]

Exemple sur SELF Anglais⁴

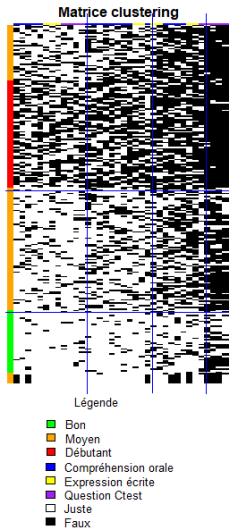


⁴Brault et al. [2020]

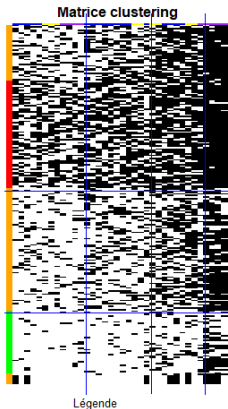
Exemple sur SELF Anglais⁴

⁴Brault et al. [2020]

Exemple sur SELF Anglais⁴

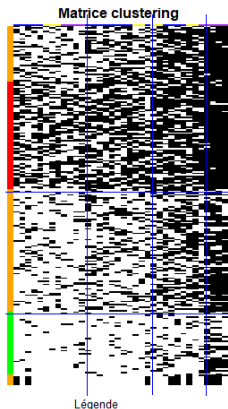


Exemple sur SELF Anglais⁴



	Bon	Moyen	Débutant	Total
Groupe Etu 1	0	36	70	106
Groupe Etu 2	0	79	0	79
Groupe Etu 3	41	2	0	43

Exemple sur SELF Anglais⁴



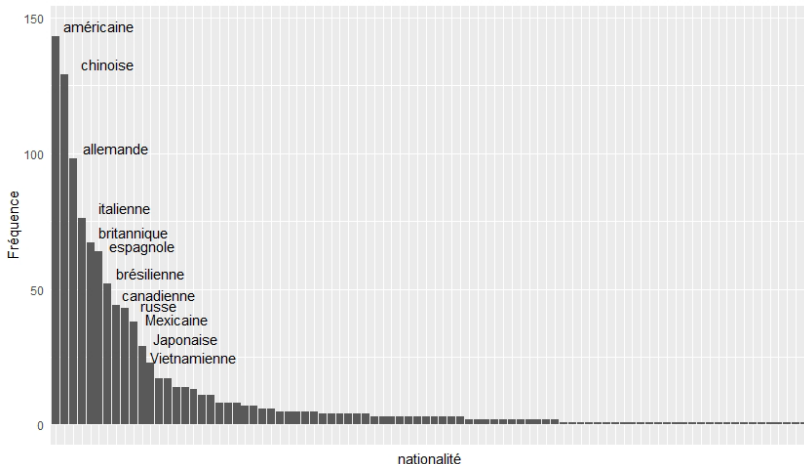
- Bon
■ Moyen
■ Débutant
■ Compréhension orale
■ Expression écrite
■ Question Ctest
■ Juste
■ Faux

	Bon	Moyen	Débutant	Total
Groupe Etu 1	0	36	70	106
Groupe Etu 2	0	79	0	79
Groupe Etu 3	41	2	0	43

	CO	Ctest	EEC	Total
Groupe item 1	5	4	3	12
Groupe item 2	9	0	2	11
Groupe item 3	4	1	4	9
Groupe item 4	0	4	0	4

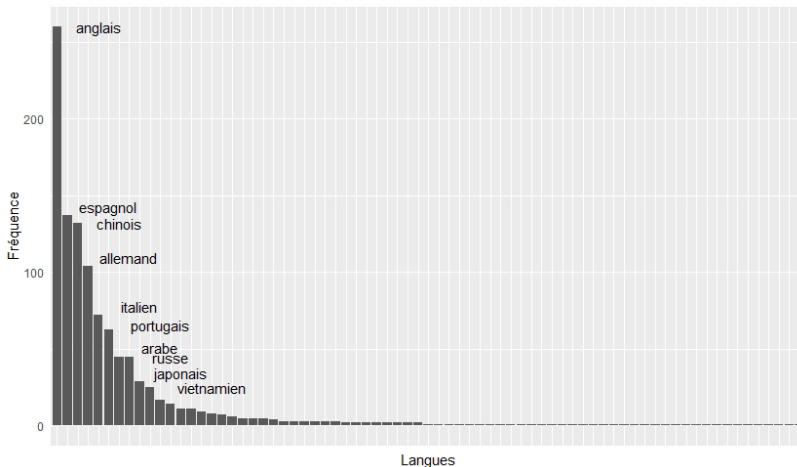
	CO	Ctest	EEC	Tout
Groupe d'item 1	52.45%	57.78%	55.66%	55.03%
Groupe d'item 2	42.24%	-	45.28%	42.8%
Groupe d'item 3	28.54%	30.19%	29.95%	29.35%
Groupe d'item 4	-	4.01%	-	4.01%
Global	42.03%	30.82%	41.93%	39.2%

Co-variables dans SELF FLE⁵



⁵Leroy [2020]

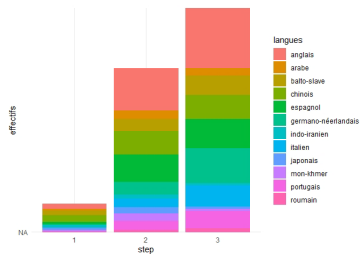
Co-variables dans SELF FLE⁵



⁵Leroy [2020]

Co-variables dans SELF FLE⁵

	Step 1	Step 2	Step 3
anglais	12	100	142
espagnol	5	64	68
chinois	17	55	57
germano-néerlandais	1	30	82
balto-slave	10	28	46
italien	0	20	50
portugais	2	21	40
arabe	4	21	17
mon-khmer	3	18	5
japonais	5	14	6
indo-iranien	7	10	5
roumain	1	6	10



⁵Leroy [2020]

Quelques problématiques

- Étude des questions non discriminantes⁶.
- Utiliser de classifications existantes⁷.
- Introduction de co-variables comme la langue maternelle⁸.
- Classification de plusieurs promotions utilisant le même test⁸.

⁶Leroy [2020]

⁷16h00 - **Marie Durand** : Clustering semi supervisé appliqué aux résultats d'apprenants débutants du polonais

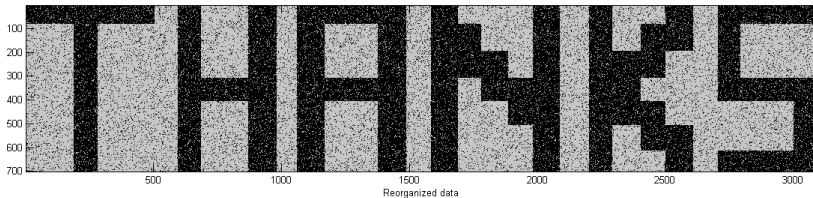
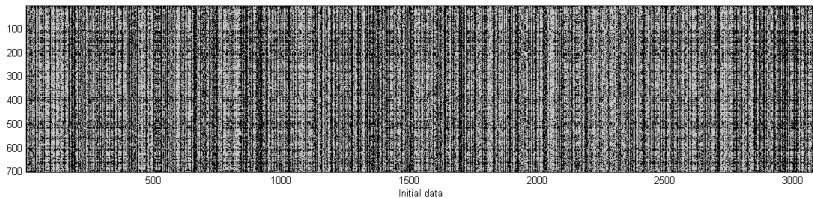
⁸13h30 - **Sophie Donnet** : Inférence de données de réseaux multipartites.

Plan

- 1 Introduction
- 2 Collaborations : exemple du test SELF
- 3 Approche par classification
- 4 Programme de la journée

- 11h00 Sébastien Georges** : Le tout est plus que la somme des parties
- 13h30 Sophie Donnet** : Inférence de données de réseaux multipartites.
- 14h30 Dominique Casanova** : Comment veiller à l'intégrité des résultats aux tests au moyen de l'analyse des données ?
- 16h00 Marie Durand** : Clustering semi supervisé appliqué aux résultats d'apprenants débutants du polonais

Merci pour votre attention et bon workshop



Bibliographie

- V. Brault, S. Coulange, M.-P. Jouannaud, F. Letué, M.-J. Martinez, and A.-C. Perret. Comment former des groupes d'étudiants homogènes à partir des résultats de self ? présentation d'un outil d'aide à la décision pour la création de groupes. In EVAL19 - Évaluation des acquisitions langagières : du formatif au certificatif, Grenoble, France, 2020.
- G. Govaert and M. Nadif. Clustering with block mixture models. Pattern Recognition, 36(2): 463–473, 2003.
- M. Leroy. Classification croisée et modèle des blocs latents. Technical report, Université Grenoble Alpes, 2020.