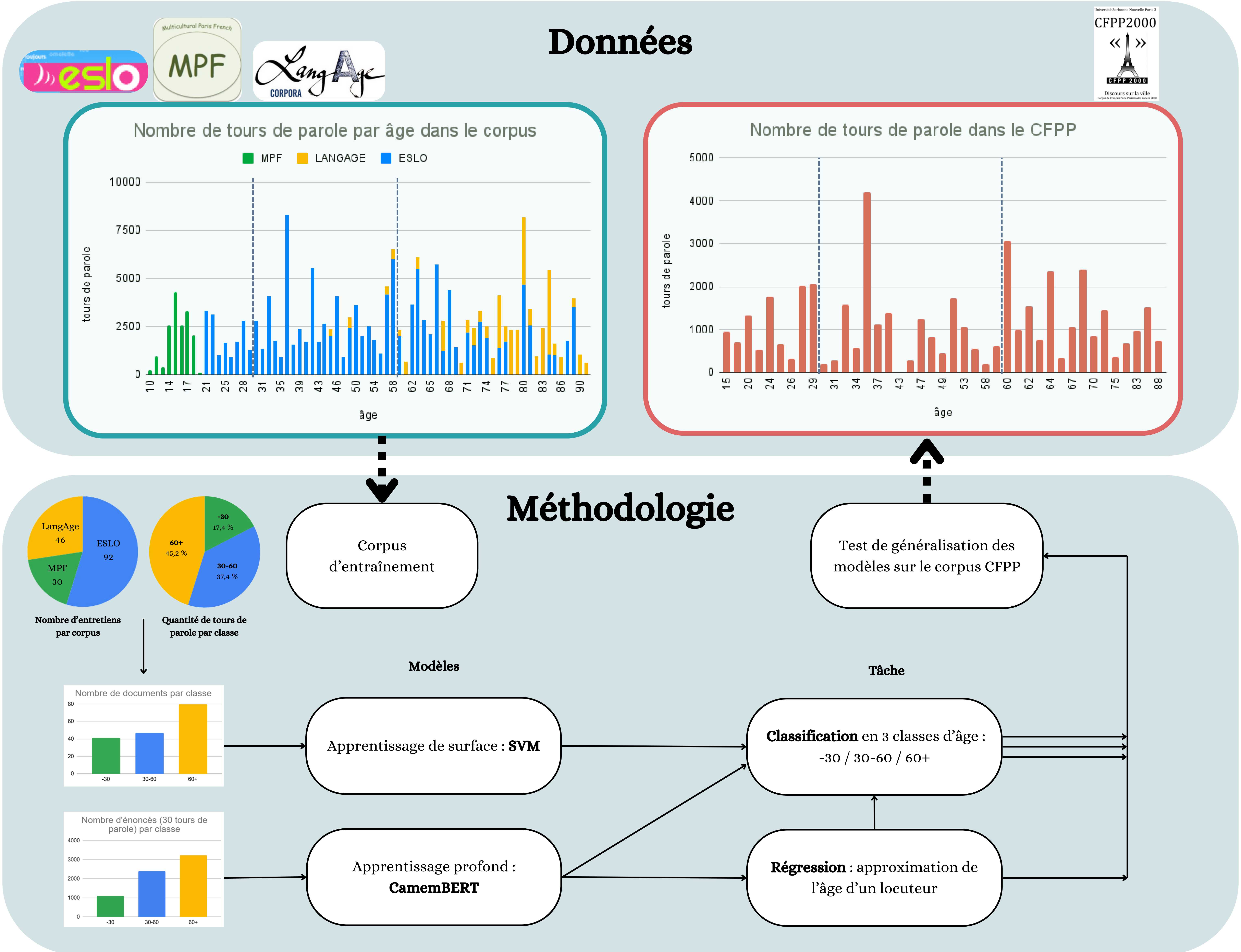


Iris Eshkol-Taravella, Angèle Barbedette, Vanessa Gaudray Bouju, Menel Mahamdi

L'âge constitue un des enjeux principaux de la sociolinguistique contemporaine en tant que facteur de variation du langage (Wagner, 2012).

Les usages individuels se stabilisent au début de l'âge adulte (Labov, 1966). Deux types de langages sont souvent étudiés :

- langage des jeunes : souvent considéré moins "standard" que celui des plus âgés (Gerstenberg & Voeste, 2015) ;
- langage des personnes âgées : marqueurs dus à des capacités cognitives et physiologiques diminuées (Kemper et al., 1992 ; Stine-Morrow & Payne, 2016).



Résultats

Modèle	Tâche	Fenêtre	Paramètres	Résultats (dataset test)			Résultats généralisation CFPP		
				Précision	Rappel	F-mesure	Précision	Rappel	F-mesure
SVM	Classification (-30 / 30-60 / 60+)	Document entier	TF-IDF unigrammes et bigrammes	0.88	0.87	0.87	0.68	0.62	0.63
CamemBERT	Classification (-30 / 30-60 / 60+)	30 tours de parole	- 8 epochs - maximum de 200 tokens - taille du batch : 12 - taux d'apprentissage : 2e-5	0.84	0.82	0.82	0.41	0.40	0.40
CamemBERT	Régression convertie en classification (-30 / 30-60 / 60+)			0.90	0.88	0.89	0.51	0.68	0.53
	Régression			Exactitude avec marge d'erreur de 10 ans : 65,6 %			Exactitude avec marge d'erreur de 10 ans : 56.1 %		

Perspectives

- Prise en compte des critères sociologiques comme facteurs de variation du langage (**catégories socio-professionnelles, niveau d'étude**) ;
- Intégration des **données prosodiques** ;
- Prise en compte des **thématiques abordées** par les locuteurs.