

La base de productions orales FRANCORAL et l'étude de l'interlangue des apprenants macédonophones de FLE : bilan et perspectives

Milena Kasaposka-Chadlovskaja

Département de langues et littératures
romanes

Faculté de philologie « Blaze Koneski »
Skopje, République de Macédoine du Nord
milena.kasaposka@flf.ukim.edu.mk

Milena Nedevska

Département de langues et littératures
romanes

Faculté de philologie « Blaze Koneski »
Skopje, République de Macédoine du Nord
milena.nedevska@flf.ukim.edu.mk

Anita Kuzmanoska

Département de langues et littératures
romanes

Faculté de philologie « Blaze Koneski »
Skopje, République de Macédoine du Nord
anitakuzmanoska@flf.ukim.edu.mk

Résumé — FRANCORAL constitue la première base électronique de productions orales d'apprenants macédonophones du français langue étrangère (FLE). Elle comprend 132 enregistrements réalisés en laboratoire dans des conditions contrôlées. La base a pour objectif de documenter l'interlangue de ces apprenants et de fournir du matériel pour analyser les spécificités de leur acquisition du français en contexte universitaire. Les données recueillies offrent ainsi un support empirique précieux pour la recherche en phonétique, morphosyntaxe, pragmatique et didactique du FLE. Cette base contribuerait de ce fait à renforcer la recherche francophone et à encourager le développement de coopérations scientifiques internationales.

Mots-clés — *base Francoral, apprenants macédonophones, FLE, interlangue, corpus oraux.*

I. INTRODUCTION

Les recherches linguistiques contemporaines s'appuient largement sur les corpus qui fournissent des données empiriques indispensables à l'analyse des phénomènes langagiers en contexte authentique. En particulier, les corpus oraux représentent une ressource majeure, car ils permettent d'observer l'usage spontané de la langue en situation réelle et facilitent l'étude du discours oral. Recueillies de manière systématique puis transcrites et annotées, ces données offrent aux chercheurs la possibilité d'examiner, dans leur contexte naturel, diverses dimensions linguistiques telles que la pragmatique, la phonétique, la syntaxe ou encore la sociolinguistique.

La langue orale est étroitement liée à son contexte d'énonciation. Contrairement à la langue écrite, elle se caractérise par sa spontanéité, son intonation, ses pauses et des structures grammaticales souvent non conventionnelles, des traits généralement négligés dans les études fondées sur des textes écrits. Les corpus oraux permettent de pallier cette lacune en saisissant les spécificités de la communication spontanée et en offrant une compréhension plus subtile de la langue en usage réel. Les corpus interlangues oraux, quant à eux, permettent de suivre le processus d'acquisition du français comme langue étrangère et constituent des outils indispensables tant pour la recherche linguistique que pour la didactique du FLE. Leur étude offre la possibilité d'identifier et de décrire les phénomènes caractéristiques de l'interlangue. Plusieurs projets universitaires et interuniversitaires francophones existent dans ce domaine. Un des projets les plus importants est le corpus *Interphonologie du français contemporain* (IPFC), créé par l'Université de Genève et deux universités japonaises (l'Université Waseda et l'Université des langues étrangères de Tokyo) et lancé en 2008 comme sous-corpus du Corpus phonologique du français contemporain. Il rassemble des apprenants de français langue étrangère ou seconde issus de diverses langues maternelles (européennes,

asiatiques et moyen-orientales). Ce corpus inclut principalement des étudiants, mais aussi des lycéens et doctorants. *Le Corpus InterLangue* (CIL), un projet mené à l'Université de Rennes 2, est un corpus franco-anglais réunissant des apprenants de français et des francophones apprenant l'anglais, qui accomplissent les mêmes tâches : *entretien, lecture à voix haute, production écrite*. *Le Corpus Didactique, Expression Orale* (DOC-STL), créé à l'Université de Lille rassemble des interactions orales en français, chinois, espagnol, portugais et italien. Le corpus *PAROLE* (PARallèle Oral en Langue Étrangère), élaboré à l'Université de Savoie en collaboration avec l'Université de Lyon rassemble les productions orales d'apprenants de l'anglais, du français et de l'italien, ainsi que des productions de locuteurs natifs, visant à analyser différents niveaux de compétence en L2.

En République de Macédoine du Nord, peu de bases électroniques sont disponibles pour documenter l'interlangue des apprenants de langues étrangères. Si des corpus ont déjà été élaborés pour l'anglais et l'allemand, aucun projet similaire n'avait encore été mené pour le français. C'est dans ce contexte que s'inscrit FRANCORAL, première initiative visant à constituer une base de productions orales en français langue étrangère recueillies auprès d'étudiants macédoniens afin de documenter et d'analyser leur acquisition du français dans un cadre universitaire. La base rassemble des enregistrements réalisés en laboratoire et fournit du matériel linguistique authentique destiné à soutenir des recherches quantitatives et qualitatives sur l'acquisition du français.

II. LA METHODOLOGIE DE CONSTITUTION DE FRANCORAL

La constitution de FRANCORAL s'est déployée en plusieurs étapes allant de la sensibilisation et de l'information des étudiants participants jusqu'à la mise en ligne finale de la base. Les stimuli élaborés incluent différentes tâches permettant de susciter des productions variées et représentatives. Les enregistrements ont été effectués en laboratoire, dans des conditions techniques contrôlées, puis soumis à un traitement rigoureux comprenant la préparation des fichiers audio, leur transcription et leur classification. Entre mi-mai et fin juin 2024, 132 productions orales ont été enregistrées. Enfin, la base a été mise en ligne sur le site de la Faculté de philologie (<https://flf.ukim.mk/francoral/>), et sur le site de l'Académie macédonienne des Sciences et des arts (*Francoral Project* (manu.edu.mk)) avec un accès libre et gratuit, ce qui garantit une diffusion large et un potentiel d'utilisation accru dans la communauté scientifique.

Pour favoriser l'émergence du discours oral en français, quatre stimuli ont été retenus : *la lecture à voix haute, le jeu de rôle, le monologue suivi et la description d'une photographie*, chacun mettant en valeur des compétences spécifiques. La lecture à voix haute permet d'analyser la ponctuation, l'intonation, la fluidité, la prononciation et la compréhension. Les textes, tirés du roman *Le Petit Prince*, s'adaptent aux niveaux A2 à B2. Ce choix se justifie par un langage accessible, des thèmes universels et une richesse narrative et dialoguée, idéale pour travailler intonations et expressions. Le jeu de rôle met l'accent sur l'interactivité. En simulant des contextes réels de communication, il permet d'évaluer les compétences interactives, sociolinguistiques et pragmatiques, ainsi que la capacité des étudiants à gérer un échange, prendre des initiatives et répondre à leurs interlocuteurs. Un autre stimulus que nous avons employé pour la réalisation de notre projet est l'exercice du « monologue suivi ». Cet exercice incite l'apprenant à développer ses idées à partir d'un document déclencheur avec la possibilité de choisir entre deux sujets. Il permet d'évaluer la structuration du discours, l'usage des connecteurs logiques, la cohérence, ainsi que l'expression personnelle et la créativité linguistique. Un dernier exercice visant à encourager la production orale est la description d'une photographie. Ce stimulus active les compétences d'observation et de description, tout en mobilisant le vocabulaire spécifique, la précision linguistique et les connaissances culturelles des apprenants.

III. BILAN DES RECHERCHES FONDEES SUR FRANCORAL

Les résultats attendus de ce projet concernent principalement l'amélioration des connaissances sur l'interlangue des apprenants macédoniens de français. L'analyse des productions permettra d'identifier les erreurs récurrentes, de mieux comprendre le processus d'acquisition et de fournir des données empiriques pour la recherche en phonétique, en morphosyntaxe, en pragmatique et en grammaire contrastive. Ces observations contribueront également à l'élaboration de programmes de FLE plus adaptés aux besoins spécifiques des étudiants macédoniens, en tenant compte des difficultés et des particularités observées dans leurs productions orales.

Jusqu'à présent, quatre recherches ont été menées sur la base de FRANCORAL. La première a présenté la démarche méthodologique adoptée pour constituer cette base, en décrivant les étapes et en soulignant les perspectives offertes pour les recherches futures sur l'interlangue des apprenants macédoniens de FLE. La deuxième a porté sur l'analyse des erreurs morphosyntaxiques et a permis de dégager les principales tendances et difficultés grammaticales observables dans leur interlangue, accentuant à la fois son instabilité et son caractère créatif. La troisième recherche a consisté en une analyse des erreurs dans l'usage des prépositions simples qui met en évidence plusieurs tendances récurrentes, notamment

des interférences directes de la langue maternelle, mais aussi certains cas de transfert. La dernière recherche s'était centrée sur les défis de la prise de parole lors de la constitution de la base présentant les vues croisées des participants et des chercheurs.

IV. ENRICHISSEMENT DE FRANCORAL ET PERSPECTIVES DE COOPERATION

Les perspectives offertes par FRANCORAL sont nombreuses. La base représente une ressource exploitable dans des domaines variés, tels que la didactique du FLE, où elle peut servir de support à la création de matériels pédagogiques adaptés, ou encore la phonologie et la prosodie, où elle permettra d'étudier les interférences entre le français et le macédonien. Du point de vue de la linguistique de corpus, FRANCORAL constitue une première étape dans la constitution d'un corpus interlangue en français pour la région. Enfin, ce projet ouvre des possibilités de coopération scientifique internationale, notamment grâce à la mise en réseau avec d'autres bases orales francophones.

Par notre participation au Symposium de la recherche scientifique francophone en Europe Centrale et Orientale (SRSF-ECO'2025), nous souhaitons donner à ce projet une visibilité accrue, notamment en Roumanie et dans l'ensemble de la région, afin de susciter des échanges et d'initier de nouvelles collaborations scientifiques.

Notre deuxième ambition est de mettre en réseau la base FRANCORAL avec d'autres corpus existants ou en cours de constitution dans les pays francophones et balkaniques. Une telle interconnexion permettrait d'enrichir considérablement les ressources disponibles pour la recherche sur l'acquisition du français, de favoriser les approches comparatives et de développer des analyses croisées entre différents contextes d'apprentissage.

Ce projet revêt un intérêt particulier pour la communauté scientifique car il contribue à combler un manque de ressources accessibles concernant les productions orales d'apprenants de français en Europe Centrale et Orientale. En intégrant les données issues de divers pays de la région, il serait possible de dresser une cartographie plus fine des profils d'apprenants, de mettre en évidence les spécificités liées aux interférences linguistiques et culturelles, et d'identifier des stratégies d'enseignement adaptées.

RÉFÉRENCES

- [1] Corpus Didactique, Expression Orale (DOC-STL) <https://www.ortolang.fr/market/corpora/doc-stl#>
- [2] Corpus InterLangue-CIL: <https://corpusinterlangue.nakala.fr/>
- [3] Corpus PAROLE: <https://talkbank.org/slabank/access/French/PAROLE.html>
- [4] Interphonologie du Français Contemporain (IPFC): <https://cblle.tufts.ac.jp/ipfc/>
- [5] Saints Cyril and Methodius University of Skopje, & Kasaposka-Chadlovská, M. (2024). Base électronique de productions orales d'apprenants macédoniens de la langue française FRANCORAL [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.14541061>