

« Du recueil de données à l'analyse des corpus en didactique de l'anglais : éthique, procédure, outils, moyens »

Pour le chercheur en didactique de l'anglais, le travail sur corpus est particulièrement enrichissant pour saisir la nature et les spécificités des productions et interactions dans la classe de langue. Le travail sur échantillon constitue une voie envisageable pour nourrir la réflexion autour de questions de recherche données. Pour ce faire, il est recommandé que la taille des échantillons d'appui soit « représentative »¹, pour satisfaire aux exigences de robustesse, validité et normes académiques habituelles de la sphère scientifique : enregistrer quelques heures d'interaction dans une salle de cours permet en effet de recueillir quelques éléments exploitables mais ne tient pas lieu d'échantillon représentatif de l'anglais parlé en situation réelle, par exemple. Les corpus peuvent donc prendre diverses formes en fonction de leur objet d'étude : des ensembles de productions écrites (scripturales ou numériques), de productions ou d'interactions orales, enregistrées en format audio ou vidéo, en situation écologique ou expérimentale. Le recueil peut alors être unique ou multiple, permettant ainsi une analyse longitudinale. Quant à la taille de ces corpus, elle peut être variable, allant de corpus écrits ou transcrits, comprenant plusieurs milliers ou centaines de milliers de mots, à des corpus plus réduits et éventuellement analysés d'après des approches plus qualitatives.

Construire un corpus s'inscrit dans une démarche scientifique globale (Dalbera, 2002) et implique une analyse spécifique en arrière-plan et le type de données que le chercheur sélectionne traduit une/des préoccupation.s sous-jacente.s précise.s, d'où la multiplicité des formes que peuvent adopter les corpus : corpus de textes, corpus d'unités de langue, ceux qui relèvent de l'écrit et ceux qui relèvent de l'oral, corpus conçus comme des échantillons de faits linguistiques, ceux qui se veulent exhaustifs dans un champ donné, etc. Dans une optique « responsable » de recherche (Narcy-Combes, 2005), il est donc essentiel pour le chercheur de penser son corpus et de le construire en fonction de ce qu'il souhaite analyser, étudier, démontrer.

Cette deuxième journée d'études organisée par l'association pour la recherche en didactique de l'anglais et acquisition (ARDAA) sera l'occasion de proposer un dialogue entre des chercheurs confirmés, des doctorants et des étudiants de Master MEEF anglais, sciences du langage et sciences de l'éducation français et internationaux, autour d'une variété de possibilités de recueil de données, de constitution de corpus, d'exploitation et d'analyse des données qu'ils contiennent.

Les échanges permettront aux jeunes chercheurs (étudiants de master et doctorants) d'identifier les principaux critères de classification de leurs corpus et leur exploitation : corpus de textes, corpus d'unités de langue, ceux qui relèvent de l'écrit et ceux qui relèvent de l'oral, corpus conçus comme des échantillons représentatifs des faits linguistiques et ceux qui se veulent exhaustifs dans un champ donné.

Les intervenants invités seront des spécialistes reconnus et incontournables des grandes notions de la journée d'étude. Ils/elles présenteront leurs corpus et les outils qu'ils/elles

¹ Par exemple, en statistiques, les échantillons strictement supérieurs à 30 individus sont considérés comme de « grands échantillons » (Couty et al, 2004 : 95).

utilisent pour analyser leurs données et répondre aux questions de recherche qu'ils.elles poursuivent.

Date : **Vendredi 24 janvier 2020**

Lieu : **Campus LSH, 23 Boulevard Albert 1^{er}, Nancy, salle G004**

Programme

9h30 – accueil, café

9h55 – Ouverture de la journée d'études par Virginie Privas-Bréauté (Nancy)

10h - 11h00 – Conférence plénière : Alex Boulton (Nancy) - *Trente ans de recherches en apprentissage sur corpus*

11h00 - 12h30 : Présentations (modérateur Elise Ouvrard, Caen)

11h - 11h45 – Clive Hamilton (Paris) - *Exploiter les corpus d'écrits universitaires : le poids de la discipline et le niveau d'expertise*

11h45 - 12h30 – Anne-Marie Voise (Paris) et Marie Potapushkina-Delfosse (Rouen) - *Corpus en classe de langues à l'école maternelle*

12h30 – 14h : pause déjeuner

Ateliers 14h – 17h30 (modérateur Joséphine Rémon, Lyon)

Atelier 1 - 14h - 14h40 – Pauline Beaupoil et Justine Paris (Paris) - *Transcrire, coder, annoter et analyser des données vidéos : tour d'horizon de quelques outils*

Atelier 2 - 14h40 - 15h15 – Atheena Johnson (Paris) - *L'impact de l'outil utilisé sur l'apprentissage*

Atelier 3 - 15h15 - 15h55 – Pascale Manoïlov, Agnès Leroux (Paris) – *Constituer et analyser un corpus d'interactions orales : d'un objectif à l'autre.*

16h00-16h15 pause- café

Atelier 4 - 16h15 – 16h50 - Stéphane Soulain (Montpellier) – *Apprendre par corps : défis méthodologiques d'une étude de transposition didactique auprès d'enseignants de langues*

Atelier 5- 16h50 – 17h30 - Margot Esnault (Nantes) - (titre à préciser)

17h30, clôture par les membres du bureau de l'ARDA

Fin à 17h45