

Lier recherche et formation professionnelle : des corpus à l'interface

Journée d'étude organisée par le l'axe CA2LI - LRL
CLARIN User Involvement Event
Vendredi 16 novembre 2018



CLARIN



MAISON DES SCIENCES DE L'HOMME
CLERMONT-FERRAND



UNIVERSITÉ
Clermont Auvergne

En sciences humaines, les recherches autour des corpus ne sont pas nouvelles mais elles ont connu dans les dix dernières années un regain d'intérêt d'une part, par le recueil de corpus non simplement écrits mais caractérisés par la multicanalité, la polysémotivité, l'(a)synchronie et l'hybridation, d'autre part, par le développement de méthodologies et d'outils de standardisation et d'annotation de ces corpus hétérogènes et enfin, par la mise en place d'infrastructures nationales et internationales permettant le partage de ces corpus dans les différentes communautés scientifiques.

Aujourd'hui, sur le plan scientifique donc, les chercheur.e.s disposent de nombreuses banques de corpus standardisés qu'ils peuvent exploiter pour leurs analyses (Ortolang, 2017). C'est vrai pour les sciences du langage qui disposent notamment aujourd'hui de corpus oraux (Childes ; TCOF, 2005 ; André et Canut, 2010 par exemple) ou de corpus écrits, depuis des corpus de textes littéraires jusqu'aux conversations synchrones ou asynchrones via blogues, tweets, SMS / textos, courriels, clavardage, forums, etc. (Corli). En sciences de l'éducation, l'on trouvera des corpus de situations d'interactions pédagogiques filmées (Muller et al., 2016 ; Falvard et Auriac-Slusarczyk, n.d. ; Veillard et Tiberghien, 2013) ou bien encore complètement médiées par les technologies (Mulce, 2015 ; Guichon et al., 2014). C'est également vrai pour les chercheur.e.s en sciences médicales ou psychologiques et psycholinguistiques qui bénéficient de corpus variés de questions de pathologies médicales (LPL, 2012 ; CORPEC, 2015 ; BiLaD, 2013-2016), qui conduisent des recherches sur des corpus spécialisés de textes médicaux (Maniez, 2011) ou encore d'entretiens cliniques (DECLICS, 2016 ; Tanguy et al., 2011 ; Vasile, 2011).

Dans la formation, initiale ou professionnelle, ces corpus sont plutôt exploités comme ressources pour décrire et montrer des phénomènes, pour les contextualiser, réaliser des études de cas ou encore pour développer une réflexion de type méta. Par exemple, les enseignant.e.s avec des apprenant.e.s, notamment de langues, utilisent des corpus le plus souvent textuels, soit constitués en banques organisées de ressources, soit rassemblés à partir de diverses sources (internet en étant la principale) sans mise en forme particulière et ce, pour le travail du lexique, de la syntaxe, pour la traduction, etc. (par exemple, Granger et al., 2002 ; Chambers, 2007 ; Boulton, 2011, 2012 ; Schaeffer-Lacroix, 2014 ; Wigham et Chanier, 2014). Dans la formation professionnelle, peuvent être proposées des tâches de réflexivité, via des activités d'auto-confrontation à sa pratique (à partir de séquences filmées) par exemple ou encore des tâches d'analyse de traces d'actions pertinentes des praticien.ne.s. (Sockett, 2014 ; Roche et Gal-Petitfaux, 2015 ; Blanes-Maestre et al., 2015 ; Sabatier et al., 2016 ; Muller et al., 2016 ; Guichon et Tellier, 2017 ; NéoPass@ction, n.d.).

On constate toutefois que le lien recherche / formation est encore trop souvent envisagé dans sa dimension de transmission des connaissances issues de la recherche sur les corpus et moins dans d'autres formes d'appropriation de ces "produits" de la recherche que constituent ces corpus par les professionnel.le.s. Si des initiatives existent, cherchant à faire le lien entre recherche et formation via ces corpus, il apparaît que l'utilisation des corpus pour la formation de praticien.ne.s, de professionnel.le.s de terrain est encore assez peu documentée.

L'objectif de la journée d'étude "Lier recherche et formation professionnelle : des corpus à l'interface" se veut exploratoire et pluridisciplinaire. Il s'agit d'établir un état des lieux des pratiques de l'utilisation des corpus notamment multimodaux en formation professionnelle - que ce soit dans le domaine de la santé, de la rééducation du langage ou de la formation des enseignants de différentes disciplines -, d'identifier les besoins qui émergent chez les formateurs/formatrices dès lors que l'on envisage l'exploitation de corpus en formation de même que les freins qui peuvent exister à l'utilisation de ces mêmes corpus.

Si la question de la définition du terme même de "corpus" se posera d'évidence, les problématiques de l'interrogation de ces corpus par les formateurs/formatrices et des tâches d'apprentissage à construire à partir de ces corpus paraissent inévitables :

- Alors qu'il est difficilement envisageable que des formateurs/formatrices se forment préalablement aux outils actuellement utilisés par les chercheur.e.s dont les corpus sont la matière première, comment des professionnel.le.s novices en matière de corpus peuvent-ils interroger les corpus pertinents pour la pratique ?
- Quelles interfaces pour interroger les corpus à des fins de formation ?
- Comment transformer des extraits de corpus en objets de formation "lisibles", accessibles pour les professionnel.le.s/praticien.ne.s de terrain ?
- Comment rendre explicite et simple les différentes dimensions du contexte dans lequel a été recueilli le corpus afin de faciliter la compréhension ?
- Quelles tâches peut-on construire à partir de ces corpus pour permettre une appropriation par les formé.e.s ?

Toutes ces questions seront envisagées dans deux grands domaines : la formation des enseignant.e.s et la formation des professionnel.le.s de santé. Elles feront l'objet de présentations de chercheurs dans une démarche privilégiant l'interaction avec le public.

Références

André, V. & Canut, E. (2010). Mise à disposition de corpus oraux interactifs : le projet TCOF (Traitement de Corpus Oraux en Français). *Pratiques*, 147-148, 35-51. En ligne <http://journals.openedition.org/pratiques/1597> ; DOI : 10.4000/pratiques.1597

Blanes-Maestre, C., Flandin, S. & Ria, L. (2015). Médiations humaines et artefactuelles en vidéoformation : le cas d'enseignants stagiaires en situation d'allo-confrontation autonomes et accompagnées. *Condition(s) enseignante(s), Conditions pour enseigner. Réalités, enjeux, défis.*, Jan 2015, Lyon, France. En ligne <hal-01350859>

Boulton, A. (2011). Bringing corpora to the masses: Free and easy tools for language learning. In N. Kübler (Ed.), *Corpora, language, teaching, and resources: From theory to practice* (pp. 69-96). Bern: Peter Lang.

Boulton, A. (2012). Beyond concordancing : multiple affordances of corpora in university language degrees. *Languages, cultures and virtual communities. Procedia : Social and Behavioral Sciences*, 34,

33-38. En ligne <http://dx.doi.org/10.1016/j.sbspro.2012.02.008>

Chambers, A. (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 3-16). Amsterdam : Rodopi.

Granger, S., Hung, J., & Petch-Tyson, S. (Dir.). (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Guichon, N. & Tellier, M. (Dir.). (2017). *Enseigner l'oral en ligne : une perspective multimodale*. Paris : Editions Didier.

Maniez, F. (2011). L'apport des corpus spécialisés en terminographie multilingue : le cas des syntagmes adjectivaux dans la langue médicale. *Les corpus et la recherche en terminologie et en traductologie, Meta*, 56(2), 391-406. En ligne <http://www.erudit.org/revue/meta/2011/v56/n2/1006183ar.html>

Muller, C., David, C. & Crozier, E. (2016). Des corpus vidéo pour la formation des enseignants de langue. *Recherches en didactique des langues et des cultures*, 13(1). En ligne <http://rdlc.revues.org/484>

Roche, L. & Gal-Petitfaux, N. (2015). Apprendre à enseigner l'Éducation Physique grâce à l'immersion sensorielle suscitée par un dispositif de vidéo-formation. *Formation et Profession*, 23(2), 49-61. En ligne <hal-01196268>

Sabatier, C., Moore, D. & Sinclair, N. (2016). Interactions et films de classe pour réfléchir la formation à l'enseignement des Mathématiques en Français langue seconde. Décrire pour mieux former ? *Recherches en didactique des langues et des cultures*, 13(1). En ligne <http://rdlc.revues.org/498> ; DOI : 10.4000/rdlc.498

Schaeffer-Lacroix, E. (2014). Utiliser des corpus numériques avec un public Lansad. *Alsic*, 17. En ligne <http://alsic.revues.org/2720> ; DOI : 10.4000/alsic.2720

Sockett, G. (2014). Corpus et perspectives pour l'enseignant : Compétences, formation, outils, besoins, activités, objectifs. *Recherches en didactique des langues et des cultures, Les Cahiers de l'Acedle*, 11(1). En ligne <http://journals.openedition.org/rdlc/1691> ; DOI : 10.4000/rdlc.1691

Tanguy, L., Fabre, C., Ho-Dac, L.-M. & Rebeyrolle, J. (2011). Caractérisation des échanges entre patients et médecins : approche outillée d'un corpus de consultations médicales. *Corpus*, 10, 137-154.

Vasile, B. (2011). Psychose et troubles du langage : aspects concourant à la constitution d'un corpus chez les patients schizophrènes. In *Corpus, données, modèles : approches qualitatives et quantitatives*, Montpellier : Praxiling. En ligne <hal-00602007>

Veillard, L., & Tiberghien, A. (2013). *Instrumentation de la recherche en Éducation. Le cas du développement d'une base de vidéos de situation d'enseignement et d'apprentissage ViSA*. Paris : Editions de la Maison des Sciences de l'Homme. DOI : 10.4000/books.editionsms.1930

Wigham, C. R. & Chanier, T. (2014). Pedagogical corpora as a means to reuse research data and analyses in teacher-training. In Colpaert, J., Aerts, A. & Oberhofer, M. (Eds). *Proceedings of the*

sixteenth International CALL Conference. Antwerp: University of Antwerp. <edutice-01018175>.

Corpus, banques de corpus, consortium et projets

BiLaD (Bilingual Language Development - Développement du langage bilingue : Enfants à développement typique et enfants avec troubles du langage), 2013-2016, ANR franco-allemand ANR-12-FRAL-0014-01

Childes (Child Language Data Exchange System) : <https://childes.talkbank.org/>

Corli (CORpus, Langues et Interactions) : <https://corli.huma-num.fr/>

CORPECS (CORPus Emergence de la communication chez l'enfant et stratégies d'étayage de l'adulte), 2015- : <https://mshl.hypotheses.org/664>

Déclics (Dispositif d'Etudes CLIniques sur les Corpus Santé), 2016- : <http://lrl.univ-bpclermont.fr/rubrique97.html> (projet ayant obtenu le soutien de la Région Auvergne Rhône Alpes)

Falvard, T. & Auriac-Slusarczyk, E. (Dir.) (n.d.) *La Base Philosophèmes*. [<http://philosophemes.univ-bpclermont.fr/>]

Guichon, N., Blin, F., Wigham, C.R. & Thouèsny, S. (2014). *ISMAEL LEarning and TEaching Corpus*. Dublin, Ireland : Centre for Translation and Textual Studies & Lyon, France : Laboratoire Interactions, Corpus, Apprentissages & Représentations.

LPL (Laboratoire parole et langage) (2012). *DySpoLec*. ORTOLANG (Open Resources and TOols for LANGuage) [www.ortolang.fr, <https://hdl.handle.net/11403/sldr000782/v1>].

Mulce (2015). *Mulce LEarning and TEaching Corpora (LETEC) Repository*. Université Clermont Auvergne : [Mulce.org](http://repository.mulce.org). [<http://repository.mulce.org>].

NéoPass@action (n.d.). *NéoPass@action* - Plateforme de ressources réalisées à partir de travaux de recherche fondés sur l'observation du travail des enseignants [<http://neo.ens-lyon.fr/neo>].

Ortolang (2017). *Banque de corpus écrits, oraux et multimodaux*. [<https://www.ortolang.fr/market/corpora>].

TCOF (Traitement de Corpus Oraux en Français), 2005 : <http://www.cnrtl.fr/corpus/tcof/>

Comité scientifique et d'organisation :

Damien Chabanal
Anne-Laure Foucher
Véronique Quanquin
Christine Rodrigues Blanchard
Ciara R. Wigham

Informations :

Date-limite pour l'inscription (gratuite) : **31 octobre 2018**

Toutes les informations concernant la journée seront mises à jour régulièrement sur le site :
<https://corefo2018.sciencesconf.org>

Contact : corefo2018@sciencesconf.org