Boulton Alex et Tyne Henry, *Des documents authentiques aux corpus: d'émarches pour l'apprentissage des langues*. (Collection Langues & didactique.) Paris: Didier, 2014, 309 pp. 978 2 278 07615 4 (broché)
doi:[10.1017/S0959269514000192](https://doi.org/10.1017/S0959269514000192)

Comme son sous-titre l'indique, l'ouvrage d'Alex Boulton et Henry Tyne s'adresse aux professeurs de langues étrangères dans leur ensemble plutôt qu'aux enseignants et didacticiens de français. Ceux-ci y trouveront cependant une mine d'informations grâce à laquelle ils pourront examiner et/ou renouveler leurs pratiques pédagogiques, qui ne sont peut-être pas aussi à jour qu'ils le croyaient. On aimerait pouvoir affirmer que, de nos jours, il reste très peu d'enseignants enfermés dans le 'paradigme traditionnel' (77), celui qui érige le professeur en guide et qui réduit l'apprenant à une 'éponge' à qui on montre ce qu'il doit apprendre et comment, et qui n'a qu'à écouter et à faire ce qu'on lui dit. La réalité risque toutefois d'être différente: la pratique qui vient d'être décrite est encore 'plus ou moins courante dans de nombreuses salles de classe' (78).

Même ceux qui se croient à jour ne le sont pas nécessairement autant qu'ils le souhaiteraient. L'évolution continue des TIC (technologies de l'information et de la communication) a rendu possible le développement de nouveaux modes d'apprentissage qui ne cessent d'évoluer. Parmi eux figure ce que les auteurs appellent l'ASC (apprentissage sur corpus ou, en anglais, *data-driven learning*), procédure – ou plutôt ensemble de procédures – où, le plus souvent, les apprenants ont le statut de bénéficiaires ('les enseignants ou concepteurs de manuels, syllabus, dictionnaires, etc. travaillent sur des corpus de données produites généralement par des locuteurs natifs de la langue cible afin de créer des supports pédagogiques', 7) ou bien d'utilisateurs ('les apprenants travaillent sur les corpus mêmes, observent les données, induisent les règles', *ibid.*). Il y a une troisième possibilité, largement laissée de côté dans ce livre, où, plutôt que d'être des bénéficiaires ou des utilisateurs, les apprenants sont à la source des données.

Deux chapitres sont consacrés au contexte dans lequel l'ASC a pris son envol. Le premier (9–40), qui s'occupe avant tout du A dans ASC, rappelle qu'une condition essentielle de tout apprentissage, autrefois ignorée mais désormais incontournable, est l'exposition à la langue cible dans tout ce que celle-ci a d'authentique. À cet égard, les TIC peuvent jouer un rôle primordial. L'authenticité des données qu'elles mettent à la disposition de l'apprenant devra cependant être gérée de façon appropriée, et c'est sur ce point que l'apprenant pourra bénéficier d'un encadrement institutionnel adapté. Le deuxième chapitre (41–79) présente les corpus (le C dans ASC) comme nouveau type de support TIC; il décrit également comment les corpus ont fait leur apparition en linguistique, non seulement pour l'étude de la langue écrite mais aussi pour celle de la langue parlée, et comment ils se sont ensuite imposés dans le cadre de la didactique des langues.

Les chapitres trois à sept s'attardent à l'ASC lui-même, à commencer par le troisième (81–119), qui part à la découverte des corpus et des techniques sur corpus. Les auteurs présentent les caractéristiques essentielles d'un corpus et d'un concordancier et montrent comment on peut passer d'une recherche basique (les occurrences d'un mot donné) à des recherches plus sophistiquées (où le contexte des occurrences est pris en compte).

Un certain nombre de considérations méthodologiques pour l'ASC sont évoquées au cours du chapitre quatre (120–143). Parmi les questions abordées se trouvent celles de la taille du corpus, de la représentativité des

données, de la variation. Plusieurs remarques sont également consacrées à la différence entre approches quantitatives et approches qualitatives.

Le chapitre cinq (144–192) offre un bilan et une analyse critique des études sur l'ASC. C'est dans ce chapitre, plus que dans tout autre, qu'on se rend compte de l'étendue des lectures des auteurs. La bibliographie en fin de volume (257–294) en fournit la preuve la plus éclatante: elle réunit quelque 460 renvois, dont un grand nombre restent méconnus dans l'espace francophone. Seul bémol: sauf erreur, toutes les études citées sont en anglais ou en français, alors qu'il n'y a pas de doute que d'importants travaux ont également été réalisés dans d'autres langues.

Restent, pour compléter le noyau de l'ouvrage, les chapitres six (193–224) et sept (225–248). Consacrées à des ressources spécifiques (des méga-corpus accessibles sur le web) et à la question des corpus sur mesure, ceux que l'on constitue soi-même (les soi-disant 'corpus de fortune'), mais aussi le web dans son ensemble, ils peuvent être lus d'une façon plus sélective, en fonction des intérêts du lecteur. Celui-ci trouvera également, dans ces deux chapitres, un aperçu de plusieurs outils d'exploitation de corpus destinés à l'apprentissage.

L'ouvrage se termine sur une conclusion générale (249–250), un index (251–256), une bibliographie impressionnante (cf. plus haut) et deux annexes, dont la première (295–298) présente, en ordre alphabétique, la liste des corpus, des outils et des ressources en ligne mentionnés tout au long du livre, la deuxième (299–309) énumérant, sous forme d'un tableau chronologique, une bonne centaine d'études empiriques parues entre 1989 et 2012, avec des détails sur le pays où elles ont été réalisées, la langue première majoritaire des apprenants, la langue étrangère étudiée, le niveau de connaissances, le contexte, le nombre de participants, etc. *Des documents authentiques aux corpus* est un guide qu'on ne pourra pas ignorer, tant il est riche en informations, conseils, suggestions . . . et mises en garde. Puisse-t-il se vendre en suffisamment d'exemplaires pour faire l'objet d'autres éditions, mises à jour au fur et à mesure.

Bert Peeters
School of Languages and Linguistics
Griffith University
170 Kessels Road
Nathan Qld 4122
Australia
b.peeters@griffith.edu.au