Convergences théoriques et méthodologiques entre la Linguistique de Corpus et l'Analyse Critique du Discours

Petru Ioan MARIAN
Université "Ştefan cel Mare" Suceava
marian_petru@yahoo.com

Abstract: This paper is focused on the use of corpora throughout discourse analysis. Critical Discourse Analysis is a socio-cultural approach of discourse which studies the general characteristics of speech/discourse communities in relation to issues such as power and gender. Similarly, Corpus Linguistics can uncover ideologies and evidence of disadvantage.
Key-words: Critical Discourse Analysis, Corpus Linguistics, power, ideology.

En se rapportant à l'Analyse Critique du Discours, Paul Baker notait: «La Prédominance des interprétations subjectives soulève des questions sur la représentativité, la sélection, la partialité, le préjugé, l'opinion. Le chercheur peut-il parler au nom du consommateur moyen des textes?»[1] (n.t.).

L'Analyse Critique du Discours est une approche interdisciplinaire du domaine théorique de l'analyse du discours qui comprend la langue comme une pratique sociale et qui suit les voies par lesquelles la domination sociale et politique est produite et reproduite par le discours. L'un des objectifs centraux des re-

[1] Paul Baker, *Using Corpora in Discourse Analysis*, Continuum, Londres, New York, 2006, p.10.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)

cherches de l'Analyse Critique du Discours est l'identification de ces zones du langage qui possèdent l'un des plus hauts potentiels significatifs pour l'étude des liens entre l'utilisation du langage et le fonctionnement des relations inégales de pouvoir.

Dans une tentative de réponse aux contestations provoquées par l'Analyse Critique du Discours, il serait opportun de la connecter à la Linguistique de corpus, une orientation méthodologique vouée à apporter un plus de rigueur à la recherche. «Enregistrement du trajet des mots, des syntagmes, des phrases, des textes et/ou des discours à un moment donné (…) et en étroite liaison avec une hypothèse de travail»[2] ou «étude du langage appuyée sur des exemples de l'utilisation de la langue dans la vie réelle»[3] *(notre traduction)*, la Linguistique de corpus est l'une des plus récentes approches linguistiques, ayant comme objectifs l'étude des mécanismes de la langue d'usage et l'identification des matrices linguistiques employées dans les discours.

Ayant des liaisons de substance avec les démarches scientifiques expérimentales positivistes qui mettent l'accent sur la mesure et la classification pour obtenir la vérité, la Linguistique de corpus, par son caractère exploratoire, est plus susceptible à générer une recherche objective que la linguistique introspective ou les orientations poststructuralistes et déconstructivistes dont l'Analyse Critique du Discours est plus proche.

Dans notre opinion, la Linguistique de corpus pourrait être cet élément méthodologique de rigueur capable de réduire les préjugés de la recherche, en rendant ainsi acceptable académiquement une approche comme l'Analyse Critique du Discours qui s'est souvent autodéfinie comme «une perspective critique de recherche: une analyse du discours avec *attitude*»[4]. (n.t.)

[2] Sanda-Maria Ardeleanu, «Les Corpus entre diversité et adéquation», in Sanda-Maria Ardeleanu *et al*., *Perspectives discursives: concepts et corpus*, Casa Editorială Demiurg, Iaşi, 2007, p. 202.

[3] Paul Baker, *op. cit.*, p.1.

[4] Teun A. van Dijk, "Multidisciplinary ACD: a plea for diversity", în *Methods of Critical Discourse Analysis*, edité par Ruth Wodak et Michael Mayer, Sage, Londres, 2001, p. 96.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)

La linguistique de corpus pratique une analyse majoritairement quantitative des traits manifestes des messages, en fournissant des informations sur la quantité, la fréquence et les relations qui s'établissent entre les phénomènes linguistiques suivis, crayonnant les concepts théoriques sur la base des observations de la langue.

Ce qui la Linguistique de corpus, préoccupée plutôt par les aspects formels de la langue, n'arrive pas à expliquer toute seule, mais l'Analyse Critique du Discours réussit, est la façon dont se construit la signification et la manière dont les influences sociopolitiques font leur place dans le discours. Puisque les sens naissent de l'interaction dynamique entre le texte, l'audience et le contexte, l'Analyse Critique du Discours fait des efforts consistants pour construire un modèle d'analyse capable de surprendre ces négociations permanentes.

De la rencontre de la linguistique de corpus avec *les recherches de l'action sociale*[5], telles que sont nommées ces formes de la recherche associées à l'examen du status-quo, qui se proposent d'intervenir et de changer, peuvent résulter des analyses intéressantes qui mettent en évidence la liaison étroite entre le langage et le pouvoir. Des orientations comme la linguistique féministe, l'analyse critique du discours, *queer theory* renvoient à la construction des identités de ceux dépourvus de pouvoir par le système de valeurs et par la voie du langage.

Examinant les applications de la méthodologie de corpus à l'analyse du discours, Douglas Biber[6] identifie *l'Analyse des pratiques sociales et des présupposés idéologiques associés à la communication* préoccupée par la construction sociale du discours. Cette orientation prend en compte le contexte social accompagnant la création et l'interprétation des textes et relègue au second plan la description linguistique des textes.

[5] Paul Baker, *op. cit.*, p. 9.

[6] Douglas Biber, Ulla Connor, Thomas A. Upton, *Discourse on the Move. Using corpus analysis to describe discourse structure*, John Benjamins Publishing Company, Amsterdam/Philadelphie, 2007, pp. 4-7.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)

Acceptant qu'entre les discours et les contextes (social, institutionnel, situationnel) où ils sont produits il y a une liaison inévitable, l'analyse du discours appuyée sur le corpus peut décrire les articulations existantes entre le milieu social et les façons d'énonciation, entre les positions idéologiques et les énoncés.

L'analyse de corpus favorise l'étude des textes politiques, des matériels didactiques, des articles de presse et contribue à l'Identification des idéologies et des stéréotypes. La plupart des choix lexicaux que l'individu opère dans le langage ne sont pas le résultat des sélections naturelles, mais de la pression des circonstances de communication et des normes sociales sur le langage.

Rayson et al. (1997) ont démontré que les sujets parlants des groupes avantagés socialement emploient plutôt des adverbes comme *actually (réellement)* et *really (vraiment)* que les groupes linguistiques moins avantagés socialement, qui, plus probablement, emploient des mots comme *say (dire)*, *said (dit)*, des numéraux et des mots tabous.[7]

Les techniques basées sur les corpus peuvent être appliquées aussi à l'étude des différences subtiles d'emploi du langage en fonction de genre: «Shalom (1997) étudie des textes publicitaires adressés aux femmes et aux hommes. Rey (2001) mène une étude sur le corpus des dialogues de la série télé *Star Trek*, analysant les différences d'emploi du langage chez les femmes et chez les hommes, tandis que Biber et Burges (2001) ont examiné la modification des différences dans les dialogues de théâtre, utilisant le corpus Archer des textes du théâtre des VII^{e} et XX^{e} siècle. (…) Sigley et Holmes (2002) ont réalisé une analyse de la fréquence et de la distribution des termes comme *girl(s)-fille(s)* et *boy(s)-garçon(s)* en cinq corpus de l'anglais britannique, en concluant que les femmes adultes sont construites linguistiquement comme immatures, accentuant leur aspect, leur dépendance, les caractères domestique et soumis. Stubbs (1996) analyse la manière dont le genre est construit dans deux discours appartenant à Robert Baden-Powell, aboutissant à la conclusion que l'idéologie peut

[7] Paul Baker, *op. cit*., p. 2.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)

habiter même derrière un mot innocent comme *heureux*. Stubbs montre que Baden-Powell (le fondateur de l'association des scouts) enseigne aux filles de rendre les autres heureux, tandis que les garçons sont instruits de vivre heureux leurs vies.»[8] *(notre traduction)*

La linguistique de corpus est nécessaire à l'Analyse Critique du Discours parce qu'elle peut fournir des données statistiques sur la fréquence d'utilisation des mots derrière lesquels se cachent des stéréotypes ou peut soutenir la création des cartes des associations du lexique, partant du principe que «l'association entre deux mots qui apparaissent souvent dans les manifestations de la langue naturelle est une épreuve meilleure pour un discours hégémonique.»[9] (n.t.). Par exemple, on peut constater que la fréquence d'association dans la même phrase des termes *despites (en dépit)* et *wheelchair (fauteuil roulant)* accroît, formant un modèle qui trahit l'horizon d'attente des individus lié à la personne invalide utilisant un fauteuil roulant.

Les études sur la fréquence d'utilisation des mots appliquées sur un corpus diachronique peuvent éclaircir l'évolution du langage, mais aussi des mentalités sur lesquelles le discours est construit: «En comparant l'emploi de la langue anglaise dans les années 1960, respectivement 1990, on a pu trouver des changements importants: le vocabulaire qui reflète l'ascension des discours capitaliste et écologiste apparaît plus fréquemment dans les années 90. En échange, des termes comme *girl (fille)* et titres comme *Mr. (Monsieur)* ou *Mrs. (Madame)* étaient plus populaires dans les années 60, ce qui suggèrerait que les discours sexistes et formels sont devenus moins répandus».[10]

«Les grands corpus offrent un appui important au projet de l'analyse du discours pour revenir aux faits authentiques de la langue. Cet appui est d'autant plus important que plusieurs études qui se proclament 'analyses du discours' opèrent avec des règles

[8] *Ibidem*, p.14.
[9] *Ibidem*, p.11.
[10] *Ibidem*, p.14.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)

abstraites partant des données inventées».[11]. Analysant les obstacles surgis dans la tentative de conversion du discours vivant, trop lié aux activités humaines domestiques, dans un objet approprié à l'étude scientifique, Teun A. van Dijk propose la solution de la reconnexion des sciences du langage à l'analyse du discours par l'intermédiaire de la linguistique de corpus.

Bibliographie

ARDELEANU, Sanda-Maria, «Les Corpus entre diversité et adéquation», in ARDELEANU, Sanda-Maria *et al*., *Perspectives discursives: concepts et corpus*, Casa Editorială Demiurg, Iaşi, 2007.

BAKER, Paul, *Using Corpora in Discourse Analysis, Continuum*, Londres, New York, 2006.

BIBER, D., CONNOR, Ulla, UPTON, Th. A*., Discourse on the Move. Using corpus analysis to describe discourse structure*, John Benjamins Publishing Company, Amsterdam/Philadelphie, 2007.

van DIJK, Teun A., *Discourse as structure and process. Discourse studies: A multidisciplinary introduction*, Sage Publications, Londres, Thousand Oaks, New Delhi, 1997.

van DIJK, Teun A., *Multidisciplinary ACD: a plea for diversity*, in *Methods of Critical Discourse Analysis*, édité par Ruth Wodak et Michael Mayer, Sage, Londres, 2001.

[11] Teun A. van Dijk, "Reconecting the science of language to discourse: Large Corpus Linguistics", în *Discourse as structure and process. Discourse studies: A multidisciplinary introduction*, Sage Publications, Londres, Thousand Oaks, New Delhi, 1997, p. 59.

BDD-A17712 © 2013 Editura Universităţii din Suceava
Provided by Diacronia.ro for IP 216.73.217.75 (2026-11-06 20:58:39 UTC)