

Revue Langage, Discours et Sociétés (LaDiSo), N°1

Médias, discours et représentations en Afrique

Numéro dirigé par Christiane Félicité EWANE

Mai 2022

PYGMIES

Cet ouvrage a été réalisé par les éditions Pygmies
contact@pygmieseditions.com
www.pygmieseditions.com
Tous droits réservés pour tous pays.

Dépôt légal.
© Éditions Pygmies, 2022
ISBN : 979-10-424-2417-6

Directrice de publication

Christiane Félicité ÉWANÉ

Comité scientifique

Alphonse Joseph Tonyè (Université de Yaoundé 1)
Christiane Félicité Ewane (Université de Yaoundé 1)
Christine Onguéné Essono (Université de Yaoundé 1)
Dili Palaï (Université de Maroua)
Edmond Biloa (Université de Yaoundé 1)
Edouard Ngamountsika (Université Marien Ngouabi)
Emmanuel Ngué Um (Université de Bertoua)
Gérard Marie Noumssi (Université de Yaoundé 1)
Germain Moïse Eba'a (Université de Yaoundé 1)
Gratien Antindogbe (Université de Buea)
Jacques Evouna (Université de Maroua)
Jean Benoît Tsofack (Université de Dschang)
Jean Tabi Manga (Université de Yaoundé 1)
Jules Assoumou (Université de Douala)
Julia Messina Ndibnu (Université de Yaoundé 1)
Julien Kilanga (Université d'Angers)
Ladislav Nzesse (Université d'Ebolowa)
Louis Martin Onguéné Essono (Université de Yaoundé 1)
Marie France Andeme (Université Omar Bongo)
Mwatha Musanji Ngalasso (Université de Bordeaux 3)
Paul Zang Zang (Université de Yaoundé 1)
Peter Blumenthal (Université de Cologne)
Valentin Feussi (Université d'Angers)
Venant Eloundou Eloundou (Université de Yaoundé 1)
Zachée Denis Bitjaa Kody (Université de Yaoundé 1)

Comité de lecture

Amos Kamsu (Université de Maroua)
Augustin Emmanuel Ébongué (Université de Buea)
Candice Guemdjom (Université de Ngaoundéré)
Chantal Ntsama (Université de Yaoundé 1)
Donald Vessah Ngou (Université de Yaoundé 1)
François Mbarga (Université de Yaoundé 1)

François Tabi Manga (Université de Yaoundé 1)
Julien Maxime Manifi Abouh (Université de Yaoundé 1)
Louis Hervé Ngafomo (Université de Yaoundé 1)
Mandou Faty-Myriam (Université de Douala)
Marie Désirée Sol Amougou (Université de Yaoundé 1)
Paul Fonkoua (Université de Yaoundé 1)
Pierre Essengué (Université de Buea)
Simplice Kengni (Université de Yaoundé 1)
Stéphanie Engola Amougou (Université de Yaoundé 1)
Tony Onguéné Mete (Université de Yaoundé 1)

Présentation du volume

Christiane Félicité EWANE (Université de Yaoundé 1)

Créée dans un contexte de dynamisation universitaire plus accrue de la formation et de la recherche doctorales, la revue *Langage, Discours et Sociétés* (LaDiSo), qui vient s'ajouter à plusieurs autres, a la particularité d'être abritée par l'École doctorale Arts, Langues et Cultures.

Spécialisée dans le domaine des sciences du langage (le langage verbal et le langage non verbal), la revue accorde la faveur à quatre (4) axes majeurs : l'analyse du discours suivant ses diverses orientations, l'acquisition et l'enseignement des langues, la dynamique des langues, etc.

Le premier numéro de LaDiSo intitulé *Médias, discours et représentations en Afrique* a pour ambition d'interroger l'interaction entre les médias, les discours et les représentations dans l'espace africain en proie à l'explosion médiatique. Cette perspective permet d'analyser les pratiques discursives, d'en étudier les représentations et de scruter les questionnements méthodologiques nouveaux qui s'originent de ce foisonnement médiatique.

Les réflexions contenues dans cette première livraison de la revue sont au nombre de onze (11). Elles s'articulent autour de quatre (4) axes majeurs : les questions méthodologiques liées à l'analyse du corpus numérique ; les représentations socio-identitaires ; l'analyse de discours ; la variation et les dynamiques langagières.

Le premier axe développe les aspects méthodologiques de l'exploitation des données médias, notamment celles d'Internet. L'hypothèse nodale de l'étude est que ce support médiatique constitue un espace de données pertinentes pour les travaux relevant des sciences du langage. Après avoir insisté sur les particularités des données numérisées, numériques et numérisées, Venant Eloundou Eloundou, dans sa contribution intitulée *Écosystèmes numériques, pourvoyeurs de données*

pour la recherche en sciences du langage : quelles postures méthodologiques, montre comment la linguistique de corpus peut pertinemment être rentable. À cet égard, tout en soulignant les écueils méthodologiques que soulève leur exploitation, l'auteur pose que les approches sociolinguistiques peuvent être abordées pour scruter les pratiques langagières, les représentations (socio)linguistiques et les activités discursives.

La deuxième articulation porte sur les discours et les représentations dans les médias. Elle regroupe quatre contributions. Alphonse Joseph Tonyè et Léopold Ngodji Tcheutou scrutent la modalité de communication de la Société Générale Cameroun en contexte de crise sanitaire. S'inscrivant dans une optique sociosémiotique, les auteurs analysent les posts Instagram, notamment la typologie des messages et le système énonciatif. Ils aboutissent à la conclusion que cet établissement bancaire a situé, dans la durée, les posts sur la lutte contre la covid-19, pour entrer dans une sorte de *storytelling* de la marque.

S'intéressant aux discours numérisés en contexte sénégalais et burkinabé, ILLI Jean analyse l'apport des médias dans la crise sociopolitique au Sénégal et au Burkina Faso, en exploitant un corpus emprunté aux réseaux sociaux numériques. L'auteur montre comment l'investissement des TIC a décuplé le potentiel contestataire des mouvements citoyens en facilitant la mobilisation de la jeunesse, particulièrement attachée à ces nouveaux moyens de communication. Il ressort de l'étude que les réseaux sociaux ont favorisé une « citoyenneté connectée » qui a eu pour effet de redéfinir l'ensemble des comportements sociaux jugés archaïques ou néfastes, au profit de nouvelles pratiques censées permettre à l'ensemble des citoyens de s'engager dans la sphère publique.

Alimentée par des acquis de la pragmatique du discours, la réflexion de Marie Renée Atangana intitulée *Formes expressives et référentielles du discours numérique sur la femme*, s'intéresse aux modalités de représentation de la femme à travers le réseau social numérique *FevaDjoss*. L'auteure analyse les actes de langage émis par des femmes qui interviennent sur cette plateforme avant d'embrayer sur les représentations.

La dernière contribution de cet axe analyse les stéréotypes dans deux réseaux sociaux au Cameroun. En s'inspirant de la

littérature consacrée à la violence verbale et l'expression de la haine dans les médias et réseaux sociaux, Sabeuya Betbeui interroge quelques discours émis par les internautes. À partir d'un échantillon de clichés relatifs aux pratiques sexuelles ou aux usages du français, l'auteure met en exergue les clivages observés sur les plans social ou communautaire, politique et culturel.

Le troisième axe développé dans ce numéro porte sur la variation et la dynamique du français dans les pratiques langagières numérisées. Dans cette perspective, Kengne Gaston François et Atiobou Voukeng Hermann montrent que les réseaux sociaux constituent des laboratoires d'expérimentation, de variation et d'enrichissement linguistiques. Ces réseaux, notamment *Facebook*, s'érigent en des espaces discursifs technologisés où les locuteurs « font la langue » et amènent le français à fonctionner dans « l'hétérogénéité, l'instabilité, la contextualité » (Robillard, 2005 : 146). La fonction de ce canal d'échanges se voit élargie de la simple logique communicationnelle, au ludique et au créatif. Pour le démontrer, les auteurs scrutent le fonctionnement des formules emphatiques qui se manifestent par des emprunts aux langues nationales, les néologismes, le camfranglais, etc. Ces expressions emphatiques sont porteuses de charges émotionnelles dans des situations interactives.

Dans la même optique, Akono Brice étudie les faits d'appropriation du français en francophonie, en s'intéressant particulièrement aux syntagmes verbaux figés, attestés dans quelques groupes WhatsApp au Cameroun. En prenant pour ancrage théorique la psychomécanique du langage, l'auteur analyse les processus de figement et les valeurs syntagmatiques, attestés comme des indicateurs de la dynamique de la langue française.

L'approche guillaumienne est aussi mobilisée par Mbella Ntomba Alain Blaise pour analyser les modalités d'actualisation du morphème *que* dans les discours numérisés des locuteurs africains. Sur la base d'un corpus constitué à partir de Facebook, l'auteur décrit systématiquement le cinétisme de *que* en deux moments. D'une part, les emplois de *que* dans la langue, d'autre part, leur usage dans le discours en termes de capacité de représentation des données de l'univers. Il aboutit à la conclusion

selon laquelle le signifié de puissance de ce morphème décrit un cycle allant de la représentation à la non-représentation de la réalité.

La dernière articulation du numéro valorise l'hypothèse que les médias et les réseaux sociaux peuvent permettre l'émergence des postures énonciatives et des stratégies de communication innovantes. Trois contributions examinent les composantes discursives. Njoungui Edoung Vivien et Tonyè Alphonse Joseph scrutent l'énonciation journalistique dans la presse écrite, *Cameroon Tribune* et *Mutations* en l'occurrence, à partir de deux faits d'actualité qui ont marqué la vie sociopolitique du Cameroun : l'opération épervier et l'élection présidentielle de 2018. Les auteurs s'interrogent notamment sur les stratégies de captation des lecteurs et la neutralité du journaliste dans son système d'énonciation. Au bout du compte, il appert que les journalistes adoptent des stratégies énonciatives qui leur permettent d'être des énonciateurs neutres, grâce à l'usage des marqueurs du discours rapporté. Par ailleurs, ils déploient des procédés énonciatifs susceptibles d'agir et de faire réagir le lecteur. Finalement, le discours journalistique est un discours à caractère performatif et à visées mercantiles.

L'article d'Asmae Halimi et Lahcen Ouasmi s'appesantit sur la caricature politique dans un organe de presse marocain *Hespress*. En convoquant la sémiolinguistique et l'énonciation, les auteurs analysent le processus discursif de la caricature. Finalement, cette recherche montre que la caricature politique dans le journal-support est étroitement liée au discours politique.

Alain Assomo enfin se focalise sur le discours publicitaire nourri par la crise sanitaire ambiante de deux entreprises de téléphonie mobile : *Orange Cameroun* et *Nexttel Cameroun*. Pour l'auteur, il s'agit de montrer comment ces entreprises inscrites dans un environnement concurrentiel ont pu réajuster leurs stratégies, en privilégiant exclusivement la communication à distance.

Tout compte fait, ce premier numéro de LaDiSo se distingue par la diversité et l'actualité des contributions. Au regard de l'engouement observé autour de ce coup d'essai, nul doute qu'il suscitera davantage plus d'émulation chez les jeunes chercheurs.

Écosystèmes numériques, pourvoyeurs de données pour la recherche en sciences du langage : quelles postures méthodologiques ?

Venant Eloundou Eloundou (Université de Yaoundé 1)

Introduction

Dans le champ des sciences du langage, l'on observe de plus en plus l'analyse des corpus issus des espaces numériques ayant favorisé l'émergence de formes discursives qui intéressent les chercheurs dans l'exploitation des communications numériques. La réflexion de Develotte (2012) présente succinctement l'historique, les problèmes juridiques et techniques que rencontre le chercheur qui explore les « communications numériques ». Ces formes constituent une source de données d'analyse qu'Emerit (2016) appelle « la boîte à outils linguistique » et peuvent fournir aux linguistes des corpus (Mourlhon-Dallies, Rakotonoelina et Reboul-Touré, 2004). À la suite de nombreux travaux portant sur la communication numérique en sciences du langage, cette réflexion voudrait participer à l'élaboration de l'appareillage méthodologique de l'écosystème numérique initiée par Paveau (2015, 2017) et Emerit (2016). Ainsi, en nous inspirant des travaux antérieurs, nous commencerons par présenter quelques acquis méthodologiques en lien avec l'écologie numérique, avant d'embrayer sur les perspectives d'analyse des données numériques dans certains domaines des sciences du langage.

1. L'environnement numérique : quel état de lieux dans les sciences du langage ?

Les réflexions initiées depuis quelques années par plusieurs chercheurs au sujet des données d'analyse issues des supports informatiques ont rendu possible une catégorisation. À cet égard, les recherches de Paveau et d'Emerit sont très édifiantes.

1.1. Le numérisé, le numérique et le numériké

Les données provenant des technologies de l'information et de la communication (TIC), exploitées dans le champ des sciences du langage, sont présentées sous trois formes par Paveau. Elles relèvent de ce que l'auteure appelle les « technologies discursives », c'est-à-dire « l'ensemble des processus de mise en discours de la langue dans un environnement technologique » (2015 : 3). Dès lors, il faut distinguer le numérisé, le numérique et le numériké.

Dans la conception de Paveau, le « numérisé » désigne tout élément susceptible d'entrer dans un environnement numérique. Il peut s'agir d'un document imprimé ou scanné, intégré dans un logiciel pour une exploitation automatique. Il est aussi question des documents disponibles sur des sites Internet et téléchargeables. Les modifications desdits documents sont plus ou moins évidentes et l'auteure (2015 : 5) de préciser que « la modifiabilité dépend du format plus que du support : une liseuse qui permet de lire du PDF, ne permettra pas plus de modifier qu'à un autre support. Elle permettra en revanche de l'annoter, ce qui est une autre forme de modifiabilité ».

Les documents « numériques », quant à eux, désignent ceux qui sont produits « en contexte électronique hors ligne, sur un ordinateur, un téléphone, une tablette, et il[s] possède[nt] toutes les caractéristiques de l'écriture au clavier et les fonctions apportées par les affordances du logiciel d'écriture » (Paveau, 2015 : 7). Ces documents se caractérisent par la possibilité de les mettre en ligne : c'est le cas des textes et des livres. Mais ils peuvent aussi rester hors ligne, à l'instar des livres et des textes en version papier.

Par ailleurs, ces documents résistent à « une énonciation augmentée ». Il est impossible de les modifier facilement. Dès lors qu'ils sont mis en ligne, ils peuvent être annotables (sauf en cas de verrouillage) et imprimables.

Sont enfin identifiés les documents « numérikés ». Leur particularité est qu'ils sont produits en ligne. Ils nécessitent un « lieu numérique accueillant la production de discours » (Paveau, 2015 : 7). On a ainsi des sites, des blogs, des réseaux sociaux, etc. Ces documents présentent quatre traits : (i) la « délinéarisation du

fil du discours », (ii) l'« augmentation énonciative », (iii) la « technogénéricité » et (iv) la « plurisémiotité » (2015 : 8).

(i) La « délinéarisation du fil du discours » (Paveau, 2015 : 8) signifie que le texte produit dans un lieu numérique est délinéarisé. Il peut accepter les éléments divers, les genres de discours différents. Cependant, cette délinéarisation n'est pas obligatoire, des blogs pouvant être consacrés à certains genres de discours.

(ii) L'augmentation énonciative est la possibilité de multiplier les sources énonciatives. Les approches *mainstream* (énonciateur unique) ne sont plus possibles. Il peut arriver que sur « un blog, un billet signé d'un "auteur" [soit] augmenté par les discussions » (Paveau, 2015 : 9). Dans ce cas, on a plusieurs énonciateurs. Ce qui invalide la perspective égocéphalocentrée (Brassac, 2010), au profit de l'« énonciation multiple ». On obtient en quelque sorte le principe de la polyphonie énonciative ; ce qui correspond à une communication synchrone.

(iii) La « technogénéricité » repose sur le fait que dans un lieu numérique, peuvent être construits ou produits des genres de discours. En effet, les « technodiscours » sont typologisés, comme le propose Maingueneau (2013). C'est le cas des twittératures, des demandes d'amitié, du style épistolaire, etc. qui constituent des « technogenres ».

Enfin, se distingue (iv) la « plurisémiotité » qui relève de la présence de plusieurs systèmes de signes dans un lieu numérique. On y retrouve par exemple : le langage verbal, le langage non verbal et le langage paraverbal, sans oublier les formes telles que les « vidéos », les « enregistrements sonores, etc. » (Paveau, 2015 : 9). Ce qui permet d'obtenir une panoplie de données intéressantes pour une analyse relevant des sciences du langage. Si les discours numérisés et numériques posent moins les problèmes de la source énonciative, le discours numérisé permet d'examiner la problématique de l'identité numérique.

1.2. L'identité numérique en question

Le lieu numérique étant caractérisé par l'augmentation énonciative, il est important de cerner le système d'identification des sources d'énonciation, notamment le sujet parlant. De manière triviale, les pseudonymes, les avatars et les

signatures sont considérés comme les traces de l'identité du locuteur¹. Il apparaît que ces indicateurs du locuteur sont très réducteurs, car « l'identité numérique recouvre un ensemble de phénomènes bien plus large » (Paveau, 2015 : 11). Elle définit ainsi « l'identité numérique » comme :

l'ensemble des données personnelles que nous possédons ou laissons à notre insu sur le Web, c'est-à-dire ce que nous faisons, disons, partageons, ressentons, aimons, détestons, recherchons, etc. C'est l'ensemble de nos traces numériques, des marques de présence en ligne. Si elle est toujours sémiotisée, elle n'est donc pas forcément verbale : un like, une inscription à un site, une commande de livre ou une réservation de billet, un paiement par carte, la publication de photos constituent autant d'éléments de notre présence en ligne que les discours langagièrement articulés que nous tenons (Paveau, 2015 : 11).

Cette identité numérique peut être un pseudonyme (*falsum nomen*) ou l'identité civile (*anthroponyme*). Quelle que soit l'option prise par l'énonciateur, quelle que soit la nomination, l'enjeu est d'exister. On pourra lire à ce sujet les travaux de Laugaa (1986) et de Martin (2012). Dans tous les cas, il se construit une sorte de sociabilité virtuelle (Casilli, 2010). Ce qui est important, c'est la possibilité d'avoir les traces sémiotisées de l'énonciateur. Elles sont sémiotisées parce qu'elles peuvent être analysées et sont dotées de sens. Autrement dit, elles livrent des informations qui peuvent permettre de scruter les discours émis en considérant les déterminants du locuteur. À la suite des réflexions de Paveau (2015 et 2017), Emerit (2016) développe la notion de lieu de corpus.

2. Écosystème numérique et lieu de corpus

En se fondant sur les travaux liés au corpus, à l'instar de Sinclair (1996), Gadet et Copeau (2007), Rastier et Pincemin (1999), Bommier-Pincemin, 1999) et Paveau, 2015), Emerit établit la différence entre un corpus et un lieu de corpus. Si les perspectives classiques attribuent trois qualités au corpus, à savoir :

¹ Pour plus de détails, on pourra se référer aux travaux de Laugaa (1986), Martin Marcienne (2006), etc.

l'exhaustivité, la stabilité et la circonscription (Sinclair, 1996), ces trois caractéristiques ne peuvent pas s'appliquer aux données numériques.

Par ailleurs, les ressources numériques ne sont pas à prendre comme des corpus. Au contraire, c'est à partir d'elles qu'un chercheur peut construire son corpus d'étude. Dans ces conditions, un lieu de corpus se définit comme « un environnement numérique » (Emerit, 2016 : 5), qui présente des données hétéroclites ou complexes. Ledit environnement a ses caractéristiques et ses modalités de description.

2.3.1. Les caractéristiques du lieu de corpus

Afin de marquer la différence entre un corpus et un lieu de corpus numérique, Emerit (2016) identifie six (06) critères qui lui sont spécifiques :

(i) La multimodalité² : un lieu de corpus ou un environnement numérique est complexe, voire hétérogène. Il renferme les signes de natures différentes. En d'autres termes, il est polysémiotique. On peut y retrouver des icônes, des mots, des émoticônes, des photos, des vidéos, des sons, etc.

(ii) L'ouverture : tout lieu de corpus est ouvert, c'est-à-dire qu'il accepte de nouveaux éléments. Les données qui s'y trouvent peuvent être modifiées, supprimées, augmentées et bloquées.

(iii) L'évolutivité : cette caractéristique est liée à la précédente. En effet, un lieu de corpus évolue dans ce sens que les données présentes « ne sont pas figées. Elles peuvent être modifiées de façon rétroactive à n'importe quel moment par les locuteurs ou l'environnement dans lequel elles apparaissent » (Emerit, 2016 : 6). Les ajouts des locuteurs ou des données, les modifications et les suppressions provoquent ainsi « une évolution du lieu de corpus ».

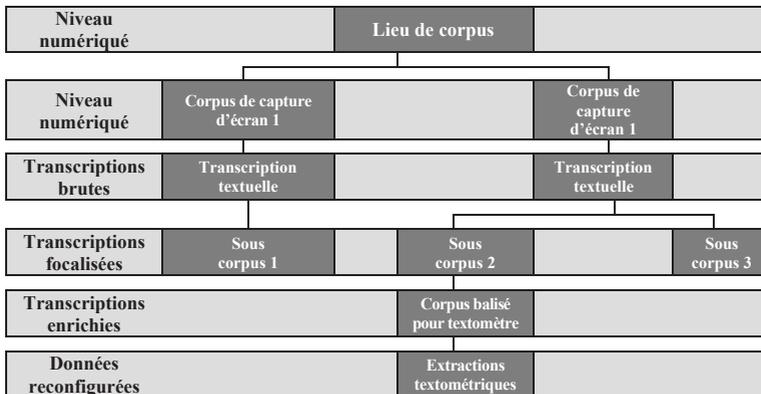
(iv) L'interactivité : selon Emerit, il arrive que des locuteurs, l'environnement numérique et le cas échéant, le chercheur qui exploite un lieu de corpus soient en interaction. Le chercheur peut stimuler des interactions entre les internautes sur un sujet précis.

² Avant Emerit, Develotte (2012 : 509) présente les corpus multimodaux en ligne comme ceux qui relèvent de la « communication à la fois multicanale et polysémiotique ». Elle (2012 : 512-513) précise que ces corpus peuvent être faiblement ou fortement multimodaux.

(v) L'interconnexion : un environnement numérique qui correspond à un lieu de corpus n'est pas une entité ou une composante isolée. Le lieu de corpus est non seulement dans un écosystème numérique, mais il entretient également des rapports avec d'autres écosystèmes numériques. On comprend qu'il peut permettre de constituer un « corpus réflexif ». Dans ces conditions, Mayaffre (2002 : 5-6) fait observer que la :

réflexivité du corpus » se fonde sur le principe selon lequel « ses constituants (articles de presse, discours politiques, pièces de théâtre ; de manière plus générale, sous-parties) renvoient les uns aux autres pour former un réseau sémantique performant dans un tout (le corpus) cohérent et autosuffisant. C'est le cas de Facebook qui peut être en relation avec Instagram, YouTube, Twitter, etc.), à partir des hyperliens.

(vi) L'idionuméricité : dans un lieu de corpus ou un écosystème numérique, il y a toujours des traces des énonciateurs ou les données personnalisées. La qualité et le nombre de ces traces dépendent de l'utilisateur ou de la technologie : c'est le cas des éléments liés au profil de l'énonciateur. Étant donné que toutes les catégories de corpus proposées par Bommier-Pincemin (1999 : 145) (corpus existant, corpus de référence, corpus de travail, corpus d'élection » ne rendent pas compte de la notion de lieu de corpus, Emerit propose une typologie adaptée à l'écosystème numérique :



Représentation arborescente du lieu de corpus (Emerit, 2016 : 10)

Le niveau 1 renvoie aux données numérisées que l'on peut consulter en ligne. Il est polymorphe, polysémotique et évolutif. Ses données sont multimodales.

Le niveau 2 correspond au corpus numérique (Paveau, 2015). Sa particularité est qu'il peut être navigable en ligne. Il est composé de données hétérogènes, chaque catégorie pouvant donner lieu à un sous corpus. Emerit (2016 : 11) pense que "chacun de ces corpus « partiels » se focalise sur un objet donné en fonction de la problématique choisie".

Le niveau 3 porte sur la segmentation ou le tri des corpus, à partir des éléments du niveau 2. À cet effet, le niveau 3 consiste à transcrire, de manière « brute », les données du niveau 2. On obtient alors un corpus textuel. À cette étape, les données textuelles transcrites ne sont pas filtrées et le corpus obtenu « peut être analysé de façon "traditionnelle" en linguistique puisqu'il est composé d'une liste d'énoncés ».

Le niveau 4 repose sur la segmentation du corpus du niveau 3. Il s'agit de répartir les données en plusieurs groupes ou « sous-catégories » (sous corpus). Le chercheur peut ainsi cibler un type de discours ou un ensemble d'énoncés qui rendent compte de sa problématique et de ses objectifs de recherche.

Emerit (2016 : 14) précise que les niveaux 3 et 4 « sont les derniers à contenir des données naturelles, bien que totalement décontextualisées, car toutes les opérations de filtrage ont été réalisées ».

Le niveau 5 se caractérise, quant à lui, par des "transcriptions enrichies. Il s'agit d'une véritable élaboration du corpus pour une exploitation informatique. En effet, l'exploitation des logiciels, dans une perspective textométrique, peut exiger des éléments supplémentaires pour une bonne utilisation. C'est le cas des balises et des métadonnées dans le texte et même le changement des formats.

Le niveau 6 renvoie aux ressources « pré-analysées de façon automatique » (Emerit, 2016 : 15). En d'autres termes, il s'agit d'un ensemble d'éléments sélectionnés et traités par un logiciel, dans le but d'obtenir « un exemplier représentatif du corpus lorsqu'il y a une grande quantité de données ». C'est aussi le lieu de procéder à leur classement.

On se rend compte que le lieu de corpus nécessite, comme le démontre Emerit (2016), l'implication de plusieurs acteurs. Chacun des acteurs a des rôles et des objectifs. Il y a d'une part, les internautes ou utilisateurs d'un « écosystème numérique » et, d'autre part, le locuteur et le chercheur, en considérant « le dialogue » entre ces acteurs. Les trois acteurs devraient avoir la même « fonction ». Dans ces conditions, « l'utilisateur est celui qui utilise les outils numériques et qui évolue dans les environnements numériques » (Emerit, 2016 : 16). Le chercheur devient un utilisateur dès lors qu'il entre en « interaction avec les autres utilisateurs de l'écosystème numérique ». Pour ce qui est du locuteur, il s'agit d'une instance énonciative qui produit des énoncés. Cependant, dans le cadre du lieu de corpus, ils 'doivent donner [au chercheur] leur consentement et le maintenir et ils peuvent participer à la recherche au travers d'interaction avec le chercheur (Emerit, 2016 : 18). Ils sont finalement des « acteurs » importants dans le champ du lieu de corpus.

Quoiqu'il en soit, le lieu de corpus relève de ce qu'Emerit appelle une « co-construction » qui fait appel aux locuteurs-utilisateurs et aux chercheurs. Ceux-ci partagent un même écosystème numérique. Si Paveau et Emerit ont réfléchi aux aspects méthodologiques des données numériques, il serait intéressant de voir comment ces ressources numériques peuvent constituer un objet d'analyse dans quelques domaines des sciences du langage.

3. Exploitation des ressources de l'écologie numérique dans le champ des sciences du langage

Les données langagières issues des écosystèmes numériques, qu'elles soient des corpus numérisés, numériques, numériqués ou du lieu de corpus, peuvent être analysées dans les disciplines des sciences du langage. Dans ces conditions, sans prétendre à l'exhaustivité, il convient de proposer quelques pistes d'analyse, notamment la sociolinguistique et l'analyse du discours.

3.1. Le champ d'analyse en sociolinguistique

Sans entrer dans le débat épistémologique concernant la définition de la sociolinguistique, il convient de reprendre ici la définition de Bulot (2013 : 6) qui stipule que la sociolinguistique « étudie la co-variance entre langue et société. Autrement dit, on cherche à comprendre les rapports dialectiques qui existent entre le changement linguistique [...] et le changement social [...] ».

Si la sociolinguistique privilégie, dans son appareillage méthodologique, le terrain social (Blanchet, 2000), désormais, il est possible d'explorer le terrain numérique qui se présente comme une variante virtuelle du terrain social. Les grands domaines élaborés par Boyer (1997 : 7) et repris par Bulot (2013 : 28) peuvent être observés dans l'écosystème numérique : les communauté(s) (sociale(s), le(s) groupes, le(s) réseaux, le(s) sujet(s), le(s) acteur(s) social/sociaux, le(s) partenaire(s) langagier(s), l'/les image(s), le/les dialectes, le/les discours, le/les texte(s), la/les pratiques ; à condition qu'ils soient réinvestis dans la communication numérique.

3.1.1. Étude des pratiques langagières

L'analyse des pratiques linguistiques ou langagières relève des usages (Calvet, 1999) de langue socialement situés. À cet égard, dans le cadre du lieu de corpus qui permet de co-construire les données et d'obtenir des corpus numérisés, un chercheur peut scruter les pratiques. En considérant que les données numérisées sont co-construites par des utilisateurs-locuteurs, il est possible de construire un corpus d'étude pertinent qui permettra d'étudier sociolinguistiquement les pratiques de langue, selon les objectifs et la problématique de l'analyse. Cependant, avec le principe de l'anonymat, il est difficile d'avoir les déterminants fiables, liés aux locuteurs-utilisateurs des environnements numériques. Les variables dépendantes qui sont fondamentales pour mieux cerner le fonctionnement des pratiques linguistiques/langagières ne sont pas faciles à obtenir ; sauf à sélectionner exclusivement les locuteurs-utilisateurs ayant utilisé des profils authentiques susceptibles d'aider à les situer socialement.

Les pratiques langagières issues du lieu de corpus peuvent relever d'une co-construction stratégiquement élaborée par le chercheur-utilisateur, avec le consentement des locuteurs-utilisateurs. Emerit (2016 : 18) fait d'ailleurs valoir que l'on doit parler de co-construction du lieu de corpus. D'un côté, les locuteurs offrent leurs conversations, leurs traces linguistiques numériques à l'étude et peuvent participer à une discussion autour de ces productions. De l'autre, le chercheur décrit, classe et analyse en s'appuyant aussi bien sur ses connaissances théoriques et méthodologiques que sur les faits de discussions qu'il entretient avec les locuteurs.

On peut même appliquer la démarche des « focus group » qui est une approche d'enquêtes sociolinguistiques. L'avantage du lieu de corpus est que le chercheur-utilisateur peut avoir plusieurs postures : le chercheur-participant, le chercheur-ami, le chercheur-utilisateur du réseau, le chercheur-espion (Emerit). Quelle que soit la posture, le principe de la co-construction s'impose.

Dans la posture du « chercheur-participant », on retrouve l'observation participante, utilisée par les sciences sociales et humaines (sociolinguistique, sociologie, anthropologie, ethnologie, etc.). Le chercheur intègre l'écosystème numérique qu'il a choisi ou construit, dans le but d'obtenir son corpus. En y participant, il conserve sa place de chercheur. S'il peut orienter stratégiquement les productions langagières, il ne saurait imposer aux locuteurs-utilisateurs les façons de parler, de voir ou de faire.

Dans sa « posture d'ami », le chercheur procède par les demandes d'amis. L'environnement Facebook permet d'élaborer une liste d'amis qui nécessite une certaine proximité sociale. Dans cette optique, pour Emerit (2016 : 21), 'les connaissances éloignées, des amis d'enfance, des collègues ou des membres de la famille peuvent tous appartenir à la catégorie des « amis Facebook », ils ne seront pas nécessairement des "amis" au sens strict du terme'. Le chercheur, dans la posture d'utilisateur du réseau, devra avoir une connaissance parfaite des réseaux en général, notamment les modes d'emploi, ses « fonctionnalités technologiques » et « ses applications, pour

entrer en relation avec les autres membres » (Emerit, 2016 : 22) de l'écosystème numérique concerné.

La posture de l'espion désigne un chercheur qui utilise un lieu de corpus sans être un locuteur-utilisateur, encore moins un ami. Qualifié de « voyeur » (Emerit, 2016 : 24), parce qu'il observe les productions discursives des locuteurs-utilisateurs, il se caractérise par sa passivité. Il peut arriver qu'il intègre un groupe d'amis sur Facebook. C'est par consentement que les locuteurs-utilisateurs acceptent son intégration. Les activités langagières qui se déroulent dans cet environnement numérique sont finalement sélectionnées et analysées par le chercheur-espion et ce dernier devient un producteur des pratiques langagières.

À cette posture de chercheur-espion, il est possible d'ajouter un espion 2. Il s'agit de celui qui se dirige vers les espaces numériques pour chercher les données de son étude. Il peut explorer les écologies numériques en fonction des possibilités qui lui sont offertes : c'est le cas d'un groupe d'amis Facebook. Cependant, la particularité de ce chercheur est que son intégration dans ces espaces ne relève pas du consentement des autres membres du groupe. L'obtention des données d'analyse devient aléatoire, il ne peut qu'explorer les écosystèmes numériques, en espérant y trouver les données pertinentes pour son étude.

3.1.2. Analyse des représentations linguistiques

Il peut arriver que certains environnements numériques soient consacrés à des questions de langue qui amènent les utilisateurs à y réfléchir. Il peut aussi s'agir des espaces d'information en ligne. Dans ce cas, une publication en lien avec les problèmes de langue peut provoquer des discours épilinguistiques/métalinguistiques (Blanchet, 2000) qui permettront au chercheur de construire son corpus d'étude. Le chercheur, par conséquent, peut adopter toutes les postures mentionnées *supra*.

Les utilisateurs auront l'occasion d'exprimer leurs perceptions et leurs attitudes à l'égard des langues concernées. Il faut préciser que ces discours, qui peuvent aussi être analysés sous l'angle des