

Entretien avec Julien LONGHI³

Carine Duteil⁴ et Christine Fèvre-Pernet⁵

Merci d'avoir accepté de répondre à nos questions pour ce numéro d'*Interfaces numériques*. Vous êtes linguiste, spécialisé en analyse du discours, sémantique et pragmatique, et vous dirigez l'*IDHN (Institut des humanités numériques, structure fédérative de l'UCP)*. Le croisement des humanités et du numérique est au cœur de votre réflexion et vous y consacrez une collection chez L'Harmattan, qui questionne notamment la médiation symbolique, sémiotique et technologique des nouveaux médias et supports. Votre dernier ouvrage (Longhi, 2018a) traite justement de la constitution de corpus numériques et de l'usage d'outils informatiques, et vous avez créé la plateforme #Idéo2017, pour l'analyse de tweets politiques. Nous allons nous entretenir avec vous sur ces aspects et nous vous invitons dans cet échange à vous appuyer sur des exemples issus de vos travaux.

Nous débiterons par une question englobante : que pensez-vous des nouvelles pratiques communicationnelles recourant au numérique ?

Je vois deux aspects intéressants dans cette formulation : la question de la nouveauté et la question du numérique. Bien sûr, les deux peuvent être liées, mais il y a, je trouve, deux dimensions distinctes dès lors que l'on s'intéresse à la communication : comment des pratiques s'adaptent à de nouveaux formats ou de nouvelles contraintes, voire comment ces pratiques modifient les formats ou leurs normes ? Et comment le numérique modifie sensiblement la communication non numérique ?

³ Julien Longhi
AGORA et Institut des humanités numériques (IDHN), Université de Cergy-Pontoise
33 Boulevard du Port, 95000 Cergy-Pontoise
julien.longhi@u-cergy.fr

⁴ Université de Limoges, carine.duteil@ensil.unilim.fr

⁵ URCA, christine.pernet@gmail.com

Dans le premier cas, on sait que le potentiel fourni par les technologies numériques est approprié par les sujets, qui en utilisent les caractéristiques voire les contraintes, pour construire des pratiques communicationnelles adaptées : c'est le cas de Twitter, qui est passé d'un simple « What are you doing now ? » à un service de *microblogging* multimédia qui intègre un grand nombre de pratiques qui jouent avec ces contraintes (par exemple l'usage du signe @ pour introduire du discours rapporté, le nombre d'échanges en lien avec les #, l'intégration d'images-textes pour dépasser la contrainte de signes, etc.).

Dans le second cas, on se focalise sur la dimension numérique de la communication, et on entre dans une distinction numérique vs non numérique. Or cette frontière devient de plus en plus poreuse, notamment dans des contextes de processus de communication généraux.

Prenons l'exemple des interviews matinales de personnalités politiques : cet exercice de communication politique, *a priori* traditionnellement éloigné d'une pratique numérique, se trouve actuellement pris dans une chaîne de communication numérique. On peut y voir une forme nouvelle d'intertextualité puisque le genre politique semble se rapprocher du genre publicitaire à travers sa dimension promotionnelle (on parle, d'ailleurs, sur les réseaux sociaux professionnels, de « *personal branding* »). Dans l'exemple ci-dessous, une personnalité annonce son passage.



Puis, il relaie ses prises de parole *via* un.e communicant.e qui met en scène son identité numérique (@benoithamon) pour lui attribuer un énoncé, soit comme citation issue d'une prise de parole (contextualisée par la photo, tweet initial), soit comme prise de parole originale reformulée (2nd tweet).

Benoît Hamon @benoithamon · 1h

En 2017, de nombreux citoyens français ont voté Macron pour éviter les politiques de Le Pen. Pourtant, sur la question migratoire par exemple, le Président a cédé aux nationalistes. Voici ce que le #PrintempsEuropéen propose pour répondre au défi migratoire 🇪🇺



2 52 88

Benoît Hamon @benoithamon [Suivre](#)

- ✓ harmonisation des systèmes d'asile européens
- ✗ fin des accords avec la Libye et la Turquie
- ✓ instauration d'un visa humanitaire
- ✗ fin des accords de Dublin

#ViveLEuropeLibre

08:39 - 13 mai 2019

35 Retweets 70 J'aime

1 35 70

Inversement, les énoncés nativement numériques, pris en charge par les identités numériques, peuvent devenir l'objet de prises de parole par les personnalités liées à ces identités numériques. Si ce cheminement est difficilement traçable, on perçoit parfois, à l'écoute d'un discours, d'une allocution, etc., des séquences discursives, une dimension « virale », proche d'éléments de langage, que l'on peut considérer comme s'intégrant dans une stratégie numérique globale (anticiper sur le fait qu'un discours va pouvoir être tweeté, ou qu'il va reprendre les séquences discursives numériques qui ont été installées dans le champ politique).

On voit que dans le cas d'une attention portée au numérique, on peut aussi saisir l'adaptation du discours aux formats : insertions d'images-textes (qui permettent d'outrepasser la limite de signes), insertion de symboles (micro, sens-interdit, etc.). Je dirais donc que les pratiques de communication liées au numérique ont une dimension créative en lien avec l'évolution des processus de communication (instantanéité, mode push, etc.) et avec l'évolution réciproque des discours et des supports numériques, qui modifient leurs contraintes et leurs caractéristiques respectives. À l'instar du discours politique qui devient un discours « promotionnel », comme souligné précédemment, on a d'ailleurs un brouillage de la frontière entre le genre politique et le genre médiatique (Longhi et Garric, 2013), ce qui nécessite de prendre en considération les conditions d'énonciation de discours oraux. Les contextes énonciatifs politiques sont parfois imbriqués dans les instances médiatiques (prenons l'exemple des émissions matinales avec des conférences de presse, des interventions dans les journaux TV) et les deux pratiques semblent se confondre, puisque le discours du média est intégré par l'instance politique à sa propre communication numérique.

Vous parlez d'identité numérique. Pouvez-vous préciser ?

Oui, on peut notamment se référer aux travaux menés en anthropologie : Julien Bonhomme⁶ explique par exemple que les anthropologues « *s'attachent à problématiser et à déconstruire l'opposition entre le "réel" et le "virtuel" à travers laquelle Internet est généralement pensé* », et qu'ils « *explorent l'univers des interactions*

⁶ Bonhomme Julien, « NUMÉRIQUE, anthropologie », *Encyclopædia Universalis*. Voir l'URL : <http://www.universalis-edu.com/encyclopedie/numerique-anthropologie/> (consulté le 6 avril 2019).

virtuelles en s'intéressant au degré de réalité et d'authenticité que leur prêtent les participants eux-mêmes ». Du point de vue de l'analyse du discours numérique, je pense que la question des mises en scène énonciatives se trouve renouvelée dans le contexte numérique : les identités ne se superposent pas, et le numérique offre une possibilité aux énonciateurs d'incarner des rôles (au sens de Ducrot, dans la définition des locuteurs/énonciateurs) et de matérialiser des identités, dont la formalisation peut aller de la simple « transposition » d'un profil sur un réseau à la constitution d'une identité narrative dotée d'un dispositif très élaboré du point de vue sémiolinguistique. Bonhomme note en outre qu'en étant parfois résumée à la présence textuelle, « *l'identité des participants se trouve dissociée des paramètres habituels qui fondent l'identité sociale dans le monde réel (genre, classe, etc.)* » : il distingue deux tendances, celle des détracteurs d'Internet qui « *s'alarment du fait que cette déconnexion permettrait toutes les tricheries et exposerait l'identité des participants à la suspicion permanente* », et celle des utopistes qui « *avançant que le libre choix de l'identité virtuelle permettrait de s'émanciper des assignations identitaires et des formes de domination en vigueur dans le monde réel* ». Il précise alors que les anthropologues « *étudient la manière dont les participants évaluent et établissent la confiance nécessaire à leurs interactions en ligne. Bien que l'identité virtuelle puisse être librement choisie et manipulée, il semble que, dans la plupart des cas, les participants s'efforcent de maintenir une certaine cohérence avec leur identité réelle* ». On voit donc que l'identité virtuelle ne se superpose pas à l'identité réelle, mais qu'elle ne s'y oppose pas forcément : c'est dans et par l'analyse des corpus que l'on peut, selon moi, observer la manière dont l'identité numérique se constitue. Dans le cadre du discours politique, il s'agit d'observer la manière dont cette identité s'articule à l'identité réelle, et d'analyser comment ces deux entités se construisent et se nourrissent réciproquement.

Plusieurs de vos travaux portent sur les mécanismes de constitution du sens en discours⁷ ; comment se construit et s'appréhende, selon vous, le sens dans ces nouveaux espaces que constituent les discours numériques ?

⁷ Notamment Longhi Julien (2008). *Objets discursifs et doxa. Essai de sémantique discursive*, Paris, L'Harmattan.

Il faut tout d'abord s'entendre sur ce que recouvre le sens, et la manière dont on peut concevoir qu'il se construit ou s'appréhende. J'inscris mes recherches, sur le plan sémantique, dans des approches à la fois textuelles, discursives et dites « dynamiques », du sens : le sens est alors considéré comme construit en corpus, à travers des configurations spécifiques, et mobilisant des niveaux de constitution divers. Cela s'inspire à la fois des travaux de François Rastier, de Pierre Cadiot et Yves-Marie Visetti, de Georges-Elia Sarfati, ou encore, d'un point de vue plus général, de postulats en lien avec la phénoménologie. On peut ainsi dire que le sens est fortement lié aux notions d'instabilité, de déploiement, de négociation et de sémiotisation. Pour en venir au numérique donc, je ne pense pas qu'il modifie la nature même de la constitution et de l'appréhension du sens, mais plutôt qu'il en déploie certaines modalités et en modifie certaines configurations. En effet, si on considère que les dynamiques sémantiques sont en particulier liées aux instances d'énonciation et aux configurations discursives, la question des identités numériques devient particulièrement centrale dans la définition de l'interaction et de la définition des rapports entre identité et altérité.

La particularité des dispositifs renouvelle aussi la mise en œuvre de certaines composantes : par exemple, avec les aspects de viralité, de référencement, de tendances, la notion d'efficacité du discours et de la stabilisation du sens se trouve corrélée à des « mesures » numériques. On peut par exemple mettre en rapport l'« efficacité » d'un tweet (nombre de retweets, de favoris) avec son thème, son lexique, ou encore son instance de production. On peut aussi utiliser des outils pour mesurer l'empreinte d'un tweet (vues, ouvertures) : bien sûr, toutes ces mesures nécessitent un regard prudent (le « succès » d'un tweet étant lié à plusieurs paramètres : audience du compte, contexte de publication, chaîne de reprises, etc.), mais ces éléments sont à prendre en considération pour appréhender de manière complète la construction et la stabilisation du sens des unités linguistiques.

Dans un article intitulé « Le tweet politique efficace comme mème textuel : du profilage à viralité » (Longhi, 2016), j'avais avancé la thèse que, dans les processus de mise en discours, il s'agit de concevoir des parcours de sens et de différenciation en lien avec une certaine mémoire discursive, ou une créativité rendue possible notamment par la brièveté des messages. Du point de vue de l'émergence du sens, « *comme les fonds sémantiques semblent des suites de points réguliers et comme les*

formes sont discrétisées par leurs points singuliers, le parcours productif ou interprétatif de ces formes et de ces fonds suppose un rythme, cellule de base de toute action » (Rastier, 2006, p.106⁸). Ce rythme prend notamment un sens nouveau au regard de la brièveté et de la viralité, ce qui permet l'émergence ou le repérage particulier des formes qui circulent. Je relie ici fond/forme au concept de profilage, qui permet à certaines formes de se distinguer, de s'individuer et, finalement, de pouvoir acquérir un statut de même : ce statut n'est pas uniquement lié à la forme elle-même (ce que nous appellerions « même linguistique », qui peut par exemple dans le discours politique, être un élément de langage qui circule, qui peut être capté, subverti, pastiché), mais à la mise en discours, dans l'espace textuel du tweet, de cette forme, ce qui nous amène à considérer des « mêmes textuels ». En effet, considérer l'opération de profilage comme centrale dans l'acquisition d'une certaine viralité pour le tweet, c'est certes tenir compte des formes qui le composent, mais c'est surtout considérer l'interaction des différentes formes présentes dans un même tweet. Pour conclure sur cette question, je dirais que le sens se construit et s'appréhende dans sa dynamique d'apparition, mais cela n'est pas pour moi différent des autres corpus ; néanmoins, le numérique joue un rôle de « catalyseur », par la circulation, la répercussion et la viralité des textes. Pour chaque type de discours numérique, il est important de caractériser les différents facteurs (genres, discours, pratiques, styles, etc.) qui contribuent à créer et stabiliser le sens (voire à le modifier, parfois *a posteriori*, par recontextualisation, dans le cadre de secondes lectures qui prennent de la distance vis-à-vis de lectures « instantanées »).

Vous analysez la communication médiée par les réseaux à travers, notamment, le projet Polititweets et plus récemment la plateforme #Idéo2017. Comment traiter les corpus issus du Web (médias sociaux, réseaux sociaux, applications...) ?

C'est une question complexe qui peut malheureusement donner lieu à des points de vue parfois caricaturaux (se pose alors la question des raisons, celles-ci étant probablement liées soit à des stratégies institutionnelles – défendre une position ; soit à la méconnaissance). Des approches qui peuvent se dire « qualitatives » (nous verrons dans la réponse à la dernière question les erreurs parfois liées à ce qualificatif)

⁸ Rastier François (2006). « Formes sémantiques et textualité ». *Langages*, n° 163, pp. 99-114.

vont privilégier des portions de corpus choisies pour la richesse d'analyse qu'elles procurent, et en proposer des analyses contextuelles ; d'autres approches peuvent procéder à des analyses plus quantitatives sur des corpus extraits *via* une API par exemple, et auront l'intérêt de pouvoir faire ressortir des traits quantitativement saillants. Chacune de ces approches a, selon moi, les qualités et les défauts qui y sont liés. Elles peuvent cependant coexister, dès lors que l'on tente d'avoir un traitement qui prenne en considération la complexité des observables.

Tout dépend notamment des questions de recherche que l'on se pose, et aussi du rapport que l'on entretient avec l'énonciation, le discours et la matérialité linguistique. Dans l'article « Tweets politiques : corrélation entre forme linguistique et information véhiculée » (Longhi, 2018d), j'avais proposé, en écho à mon travail d'HDR, une démarche de constitution de corpus qui s'inspire de la démarche de l'enquête de terrain du sociologue ou de l'anthropologue (voir Longhi, 2015, à propos de Beaud & Weber, 2012⁹). On peut chercher à aborder le corpus comme un « terrain » d'analyse.

Il convient alors de :

- se poser une question de départ ;
- passer du thème à la question d'enquête ;
- transformer « *une question "abstraite" en une série décomposée de pratiques sociales et d'événements* » (Beaud & Weber 2012, p. 36).

Si on cherche à caractériser les liens entre le statut énonciatif des tweets et l'information qu'ils véhiculent, par exemple, un « mix méthodologique » entre des moyens relevant tantôt de ce qui est qualifié d'analyse qualitative, tantôt d'analyse quantitative, semble nécessaire. Je pense que l'opposition « quantité »/« qualité » radicalise des positionnements qui pourraient être articulés (d'ailleurs bien souvent cette opposition est une mauvaise transposition des difficultés à articuler des modèles formels, théoriques, computationnels, etc. ; voir le travail de J.G. Meunier à ce sujet).

D'un côté, le « bon » corpus pourrait être qualifié ainsi parce qu'il est un « grand » corpus (et parce que la méthode se fonde sur la quantité

⁹ Beaud Stéphane et Weber Florence (2012). *Guide de l'enquête de terrain*, Paris, La Découverte.

des données, par le biais d'outils informatisés), alors que d'un autre côté, la « qualité » du corpus et des méthodes d'analyse (plus « manuelles ») pourraient prévaloir, quitte à travailler sur des corpus de moindre envergure. D'une manière générale, donc, mon objectif est d'apporter aux corpus des traitements qualitatifs à partir d'indicateurs linguistiques minutieux, appliqués à des quantités de données importantes, et d'opérer un va-et-vient entre les résultats quantifiés et leur spécification dans les observables.

Comment répondre à cette exigence ?

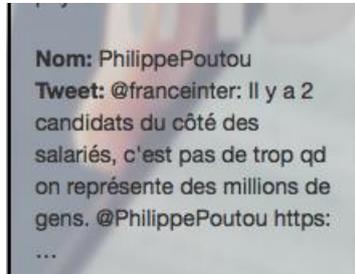
Je propose deux moyens :

- la constitution de corpus minutieusement documentés :

Dans le cadre du corpus Polititweets (Longhi *et al.*, 2014), le recours au format TEI a permis un enrichissement des données textuelles (tweets) par un grand nombre de métadonnées qui contribuent à extraire informatiquement ce qui relève de l'écologie des données (contexte spatio-temporel, interaction, etc.). Selon sa page Wikipédia (qui reprend notamment les grandes lignes du site <https://tei-c.org/>), la TEI, ou Text Encoding Initiative (initiative pour l'encodage du texte) « est une communauté académique internationale dans le champ des humanités numériques visant à définir des recommandations pour l'encodage de documents textuels ». Son modèle théorique « s'est adapté à différentes technologies, d'abord sous la forme d'une DTD SGML, puis XML », et des comités démocratiques et internationaux conduisent « la maintenance et la croissance du schéma », rédigent la documentation, développent des outils génériques, assurent le support sur des listes de diffusions (pour plus de détails, voir Longhi, 2017d, sur les enjeux institutionnels, juridiques, techniques et philologiques d'un corpus de tweets). Le corpus Polititweets a fait l'objet d'une grande attention concernant le choix des métadonnées à introduire (type de matériel, nombre de retweets, etc.) et les balises caractéristiques de certains signes spécifiques (balises de #, de @, etc.).

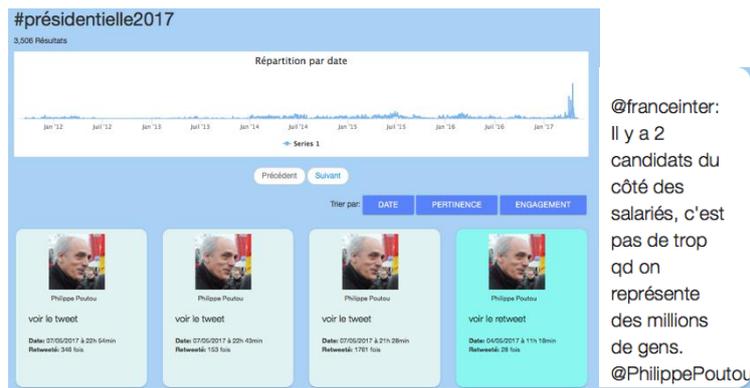
- la constitution d'interfaces articulant outils statistiques et comparaison de corpus, et navigation pour un retour aux données natives :

Depuis 2017, mes projets de recherche visent à produire des interfaces qui permettent l'articulation des deux approches. Ceci était le



Visualisation d'un message potentiellement intéressant

C'est un tweet de @franceinter retweeté par Philippe Poutou (donc un tweet circulant, pour rejoindre la thématique du volume) : il pouvait aller dans la partie « moteur de recherche ». Il pouvait alors retrouver le tweet en question.



Exploration par le moteur de recherche et obtention du résultat

L'utilisateur pouvait ensuite l'ouvrir, par un simple clic, dans l'interface de Twitter, c'est-à-dire dans son environnement natif.



Prise en compte du message dans l'environnement Twitter

On peut alors retrouver les caractéristiques des approches « contextualisantes » décrites par (Paveau¹⁰, 2013) (« *contextualisation technorelationnelle* », « *investigabilité du discours* », « *technoconversationnalité* »). On voit sur cette image le lecteur de vidéo qui propose une interview (le contenu textuel du tweet, le discours rapporté, est en fait un extrait de cette vidéo, qui est une interview de Philippe Poutou). Un va-et-vient est rendu possible entre les deux modes d'appréhension : l'analyse contextualisée nous semble ici trouver un intérêt car elle s'est focalisée sur un message qui contenait le terme « salarié », élément visiblement saillant dans le corpus de Philippe Poutou (repérage statistique et en collocations). Si l'accès au corpus se faisait uniquement par l'interface de Twitter, nous n'aurions pas d'informations sur la spécificité de ce terme pour le candidat en question, la manière dont il en fait usage, comment il se distingue de ses concurrents par le sens qu'il lui accorde. Le discours rapporté à l'ère du numérique ne doit, selon nous, pas être analysé avec des nouveaux concepts qui « augmentent » les concepts traditionnels de l'analyse du discours, mais être analysé avec les concepts traditionnels selon **une nouvelle manière de suivre les observables**. S'il est impossible d'épuiser l'analyse d'un fait de discours, qui va sans cesse être repris,

¹⁰ Paveau Marie-Anne (2013). « Analyse discursive des réseaux sociaux numériques ». *Dictionnaire d'analyse du discours numérique, Technologies discursives*, [Carnet de recherche]. Voir l'URL : <http://technodiscours.hypotheses.org/?p=431> (consulté le 15 septembre 2017).

commenté, cité, modifié (si le compte change d'avatar, si les fonctionnalités sont modifiées), il est néanmoins possible, et c'est l'enjeu de ma recherche, de fournir des moyens d'accès à ces observables, et des représentations intelligibles de ces particularités. Je parle, depuis 2018 (et dans un article en cours d'expertise), de « *corpus réfléchis* » (en écho au « *corpus réflexif* » de Mayaffre, 2002¹¹) : un corpus qui peut effectuer un « retour sur lui-même », soit dans sa matérialité formelle, soit dans son environnement natif). Je pense qu'on peut concevoir un nouveau moyen d'analyse du discours numérique à travers l'usage de corpus réfléchis et d'interfaces propices à rendre compte des différentes caractéristiques des observables.

En quoi la problématique de l'énonciation vous semble-t-elle intéressante à traiter dans ce contexte ?

Elle me semble importante à prendre en compte afin de questionner les notions de sujet, d'identité, et d'altérité notamment. Rappelons que selon Benveniste, l'énonciation est l'événement historique constitué par le fait qu'un énoncé a été produit, c'est-à-dire qu'une phrase a été réalisée. L'énonciation, qui est l'acte même de produire un énoncé, accomplit ce que Benveniste qualifie de « *conversion du langage en discours* » (Benveniste, 1966¹², p.254). Cet acte individuel d'appropriation de la langue constitue la première marque formelle de toute énonciation. Il est pris en charge par un énonciateur, dans un cadre spatio-temporel donné, et il est destiné à un co-énonciateur. Dans le contexte de discours numériques, et comme indiqué dans la question précédente sur les corpus, on peut reconstituer les conditions d'énonciation, en lien avec des métadonnées associées aux données, et on peut aussi, selon les contextes, retourner aux données natives qui rendent compte de la situation d'énonciation (échanges en ligne, supports multimédia, etc.). Certes, cette reconstitution peut parfois donner l'impression d'être un simulacre, puisqu'il y a des choix opérés en termes de balisage, d'outils, de représentation, etc. Néanmoins, il faut reconnaître que l'appréhension de l'énonciation est peut-être à considérer comme une tension à construire avec les observables, et à rendre disponible sous différentes formes et manières. Bien sûr, le

¹¹ Mayaffre Damon (2002). *Les corpus réflexifs : entre architextualité et hypertextualité*, Corpus, 1.

¹² Benveniste Emile (1966). *Problèmes de linguistique générale*, Paris, Gallimard, coll. « Bibliothèque des sciences humaines ».

« contexte » peut être reconstitué selon une appréhension plus large, avec le contexte socio-historique, politique, interactionnel. Ceci est perceptible par des marques linguistiques (topoï, mots du discours, axiologie, etc.), puisque comme l'indique Ducrot, l'énonciation laisse sa marque dans le sens de l'énoncé. Avec la question des identités numériques, il est donc très intéressant d'observer la manière dont l'énonciation s'analyse en discours, à travers les traces qui sont laissées par « l'appropriation de la langue en discours », ce qui relève des conditions d'énonciation ou ce qui relève de pratiques numériques indissociables de l'énonciation elle-même. Pour appréhender cette question, deux articles publiés dans le collectif *L'énonciation aujourd'hui, un concept clé des sciences du langage*¹³, sous la direction de Marion Colas-Blaise, Laurent Perrin et Gian Maria Tore, me semblent utiles. Dans un article intitulé « Altérité de la parole et socialité du sens : énonciation et perception d'autrui », Antonino Bondi propose de « *donner un fondement phénoménologique à la théorie sémio-linguistique* » pour penser l'activité de langage comme « *une perception, c'est-à-dire comme une activité générique de relation à, accès à (au monde), de déplacement constant du ou des sujets, d'ajustement dialogique, pragmatique et narratif selon, ou mieux, sur le fond d'un "fond" à la fois expressif et perceptif, normatif, social et institué* » (p. 383). Cette dimension perceptive est thématifiée en considérant le discours comme optimisation de l'expérience par Pierluigi Basso Fossali, pour qui « *l'énonciation pluralise donc les scénarisations des valeurs, en utilisant des médiations sémiotiques spécifiques (langages) qui garantiraient une "prise" commune entre les identités réclamées et les altérités imputées* » (p. 400). Considérée ainsi, l'énonciation, saisissable à travers les pratiques discursives, s'affirme comme un niveau d'analyse pertinent pour l'étude du discours numérique : les interfaces, les réseaux, les supports, etc., contribuent aux scénarisations dans lesquelles les identités et les altérités se construisent, et au cœur desquelles les ajustements sémio-linguistiques s'effectuent. Les traces discursives sont autant d'observables que l'on peut analyser à travers les corpus réfléchis évoqués dans la question précédente.

¹³ Colas-Blaise Marion, Perrin Laurent et Tore Gian Maria (dir.) (2016). *L'énonciation aujourd'hui, un concept clé des sciences du langage*. Limoges, Lambert-Lucas.

Pour finir, et en guise de conclusion, nous aimerions tenter de mieux circonscrire avec vous le champ des humanités numériques. Comment penser en leur sein l'interdisciplinarité ?

Les humanités numériques (HN) nécessitent une inter/pluri/trans-disciplinarité à plusieurs niveaux. Cette complexité des rapports entre les disciplines (doit-on faire coexister des disciplines, les articuler, circonscrire des tâches spécifiques aux unes et aux autres ? les projets en HN sont souvent à géométrie variable à ce sujet) est souvent le fruit de démarches liées à des contextes précis et à des objets de recherche spécifiques. Il faut déjà prendre en compte la diversité des humanités, qui ne se superposent pas forcément, du point de vue institutionnel, aux « Sciences humaines », ni aux « sciences humaines et sociales ». De l'autre côté, le numérique peut soit renvoyer à l'informatique, selon une acception parfois technique, soit aux usages, ou aux outils, sans que leur conception soit forcément prise en compte. Aussi, il me semble que cette réflexion sur les HN doit nécessairement s'accompagner, de manière aussi importante que pour les humanités, d'une définition de ce que recouvre le « numérique », et le recours plus large à l'informatique. Il faut que l'informatique rencontre les humanités, et pas seulement que les humanités utilisent l'informatique pour les besoins numériques.

Milad Doueïhi (2015, p.711¹⁴) décrit ainsi le passage de l'informatique au numérique :

« De l'informatique (qui certes n'a pas entièrement disparu) au numérique, on passe d'une technicité, souvent exagérée et cultivée pour elle-même, mais exigeant une certaine compétence technique, à des usages plus communs, exigeant d'autres compétences : celles que valorise une nouvelle sociabilité en ligne, peuplée de textes, animée par des "partages". Et aujourd'hui, c'est par rapport à cette pratique numérique populaire que les travaux en humanités numériques doivent être aussi pensés. »

Réciproquement, certaines des tâches, ou des objectifs, des humanités numériques sont abordés par les sciences informatiques en tant qu'objets de recherche, mais ne sont que rarement investies dans les projets en HN, si ce n'est sous la forme de la partie « ingénierie » des

¹⁴ Doueïhi Milad (2015). « Quelles humanités numériques ? ». *Critique*, vol. 819-820, pp. 704-711.

projets. Ceci est d'ailleurs questionné par un groupe de travail issu de la communauté des sciences de la donnée : le Groupe de travail (GT) DAHLIA (<http://dahlia.egc.asso.fr>), qui « *a pour but de réunir les acteurs (chercheurs ou institutions) qui s'intéressent, dans le cadre des humanités numériques, voire le patrimoine culturel, à la gestion des données mais aussi à leur analyse afin de produire des connaissances* ».

Avant d'entrer donc dans la pratique des humanités numériques, il y a, je crois, un socle méthodologique et épistémologique à clarifier autour, notamment, d'oppositions telles que qualitatif/quantitatif, conceptuel/formel, sémiotique/numérique. Bien souvent, certaines de ces oppositions sont utilisées l'une pour l'autre, ce qui crée des confusions. Ceci est décrit par (Meunier, 2018)¹⁵, et amène à des considérations qui changent notre rapport à l'informatique. Pour Meunier, « *un modèle formel ne porte pas nécessairement sur du quantitatif. Aussi, on trouvera des formalismes logiques, géométriques, topologiques, grammaticaux, etc. Certains utiliseront des symboles iconiques (graphes, images, etc.). Dans tous ces formalismes, on pourra trouver divers types de symboles : tels des constantes, des variables, des opérateurs, etc.* ». Ce qui peut donc distinguer une approche informatique des humanités numériques, ce ne serait pas forcément l'intérêt premier pour l'ordinateur, ou le recours à des méthodes quantitatives, mais plutôt la formalisation des connaissances. Or, « *ces modèles formels mathématiques sont omniprésents tant dans les sciences naturelles que dans les sciences humaines et sociales* » (*ibid*). Ainsi, lorsque Meunier s'intéresse à la difficulté de la notion de computationnalité pour la sémiotique, il précise qu'« *une théorie sémiotique de type computationnelle ne peut exister que si elle en appelle à de la modélisation formelle de type mathématique (non pas nécessairement quantitatif), dont les énoncés, formules ou équations permettent la calculabilité* ». L'interdisciplinarité des humanités numériques devrait donc être selon moi une interdisciplinarité « profonde », au sens où il y a une compréhension nécessaire non seulement des objets de recherche, des méthodes et des outils, mais plus largement des modèles et des manières de concevoir l'activité scientifique.

¹⁵ Meunier Jean-Guy (2018). « Vers une sémiotique computationnelle ? ». *Applied semiotics*, 26, <http://french.chass.utoronto.ca/as-sa/ASSA-No26/26-6.pdf>

Pour en proposer une définition, on peut bien sûr se référer au « Manifeste des Digital Humanities » (<http://tcp.hypotheses.org/318>) (cette définition est proposée par les « *acteurs ou observateurs des digital humanities (humanités numériques) (...) réunis à Paris lors du THATCamp des 18 et 19 mai 2010* ») :

1. Définition

1. Le tournant numérique pris par la société modifie et interroge les conditions de production et de diffusion des savoirs.

2. Pour nous, les digital humanities concernent l'ensemble des Sciences humaines et sociales, des Arts et des Lettres. Les digital humanities ne font pas table rase du passé. Elles s'appuient, au contraire, sur l'ensemble des paradigmes, savoir-faire et connaissances propres à ces disciplines, tout en mobilisant les outils et les perspectives singulières du champ du numérique.

3. Les digital humanities désignent une transdiscipline, porteuse des méthodes, des dispositifs et des perspectives heuristiques liés au numérique dans le domaine des Sciences humaines et sociales.

Relativement à ce qui a été dit précédemment, je synthétiserai donc cette définition en proposant que **les humanités numériques soient considérées comme le point de rencontre de plusieurs disciplines scientifiques, qui relèvent des humanités et de l'informatique, ayant pour objet l'étude des « données » provenant du champ des humanités (productions...), avec des méthodes, outils, ou données, liés au numérique, et proposant une construction réciproque des objets, méthodes et théories, en élaborant conjointement un modèle théorique, conceptuel, formel et computationnel, qui donne corps et cohérence à l'étude d'un objet de recherche.** L'articulation de ces différents modèles est de nature à créer une relation complémentaire (et non utilitaire) entre les différents acteurs d'un même projet. Cela a deux vertus fonctionnelles : ne pas séparer la conceptualisation de la technique (il est par exemple important que le choix de balises TEI soit passé au regard de la nature sémiotique des données, et ne soit pas vu comme une tâche technique), et ne pas déléguer des pans d'une recherche à des outils qui sont perçus comme des « boîtes noires » par certains participants.

Merci à Julien Longhi d'avoir répondu à nos questions.

Éléments de bibliographie

- Longhi Julien (2018a).** *Du discours comme champ au corpus comme terrain. Contribution méthodologique à l'analyse sémantique du discours*, Paris, L'Harmattan, coll. « Humanités numériques ».
- Longhi Julien (2018b).** « Le projet #Idéo2017 : Quelles implications du/de la chercheur-e en tant qu'acteur-trice potentiel du changement social ? Exemplification à partir du discours politique numérique ». *Cahiers de linguistique*, n°44/2, pp. 103-116.
- Longhi Julien (2018c).** « L'écriture nativement numérique, de Twitter à Youtube : pour une approche non-conversionnelle des processus créatifs ». *Le français aujourd'hui*, n° 200, pp. 43-54.
- Longhi Julien (2018d).** « Tweets politiques : corrélation entre forme linguistique et information véhiculée ». In Mercier A. et Pignard-Cheynel N. (dirs.), *#Info. Partager et commenter l'info sur Twitter et Facebook*, Paris, Editions de la Fondation MSH, pp. 295-314.
- Longhi Julien (2017a).** « Humanités, numérique : des corpus au sens, du sens aux corpus », *Questions de communication – Humanités numériques, corpus et sens*, n° 31, pp. 7-17.
- Longhi Julien (2017b).** « Linguistique de corpus appliquée ». *Etudes de linguistique appliquée*, n° 188, pp. 397-401.
- Longhi Julien (2017d).** « Le corpus Polititweets : enjeux institutionnels, juridiques, techniques et philologiques ». In Wigwam C.R. & Ledegen G. (dir.) *Corpus de communication médiée par les réseaux : construction, structuration, analyse*, Paris, L'Harmattan, coll. « Humanités Numériques », pp. 37-50.
- Longhi Julien (2016).** « Le tweet politique efficace comme mème textuel : du profilage à viralité ». *Travaux de linguistique*, n° 73, pp. 107-126.
- Longhi Julien (2015).** *La Théorie des objets discursifs : concepts, méthodes, contributions*. Habilitation à diriger des recherches, soutenue le 17 juin 2015 à l'Université de Cergy-Pontoise.
- Longhi Julien, Marinica Claudia, Borzic Boris, Alkhouli Abdulhafiz (2014). « Polititweets, corpus de tweets provenant de comptes politiques influents ». In Chanier T. (dir.), *Banque de corpus CoMeRe*. Ortolang.fr, Nancy.

Longhi Julien et Garric Nathalie (2013). « Théoriser le genre pour déjouer ses frontières et construire le sens ». *Pratiques*, vol. 157-158, pp. 31-46.

Longhi Julien (2012). « L'énonciation et les voix du discours ». *Tranel (Travaux neuchâtelois de linguistique)*, n° 56, <http://www.unine.ch/tranel/home/tous-les-numeros/tranel-56.html>.

Longhi Julien (2011). *Visées discursives et dynamiques du sens commun*, Paris, L'Harmattan, coll. « Sémantiques ».

Longhi Julien (2008). *Objets discursifs et doxa : essai de sémantique discursive*, Paris, L'Harmattan, coll. « Sémantiques ».