

VOS POUCES POUR LA SCIENCE

LOUISE-AMÉLIE COUGNON

**UN PROJET D'ÉTUDE
DE LA COMMUNICATION
« NUMÉRIQUE »**



Le projet *Vos Pouces pour la Science* est un projet de collecte chapeauté de 2016 à 2017 par l'UCL et financé par le FNRS. Il est issu d'une collaboration entre trois domaines scientifiques : la sociolinguistique, le traitement automatique du langage et la psychologie. Après un an de collecte, nous avons figé les données que nous commençons à analyser. À ce stade, nous sommes une douzaine de collaborateurs scientifiques de France et de Belgique, tous domaines confondus. Une entreprise, *Moodwalk*, nous aide aussi à gérer certains aspects techniques.

OBJECTIFS

Le projet *Vos Pouces pour la Science* vise à étudier les compétences de la population au contact des nouveaux médias de communication en général (réseaux sociaux, messages électroniques...). Contrairement aux idées reçues et véhiculées dans la presse et dans les milieux de l'éducation, nous partons de l'hypothèse que l'utilisation des nouveaux médias chez les classes d'âge de plus en plus jeunes ne mène pas à une incompétence linguistique (baisse du niveau de l'orthographe, méconnaissance des règles grammaticales...), mais plutôt à une « pluricom pétence » qui amènerait chaque locuteur à jongler avec son code à chaque changement de situation, d'interlocuteur et de médium de communication. Nous basons cette hypothèse sur les conclusions de notre recherche doctorale, menée sur la variation de l'écrit sms, et qui laissait entrevoir de telles tendances. Nous nous demandons donc dans quelle mesure ces nouvelles pratiques scripturales sont liées au « potentiel créatif » de chaque locuteur.

Pour étudier les compétences linguistiques de la population, nous devons observer les pratiques dans différentes situations de communication aux finalités diverses : pour fixer un rendez-vous, pour assurer le travail, pour communiquer en classe ou sur Facebook entre amis ou encore pour réussir une dictée. Afin de répondre à notre objectif linguistique, nous regardons comment la langue varie chez une même personne en fonction de ces diverses situations et nous analysons à quel point et à quelle fréquence cette personne fait varier sa langue et commet des « fautes

d'orthographe » (ceci comprenant autant des erreurs orthographiques que des variations ludiques sur le code). À cela, nous ajoutons l'étude d'autres paramètres qui influent sur notre production linguistique : notre psychologie et notre rapport aux autres. En effet, nous enquêtons également sur les compétences sociales et psychologiques de la population au contact de ces nouveaux médias. Nous nous posons ici aussi de nombreuses questions :

- Les nouveaux médias nous ont-ils incités à réinventer nos modes de communication ?
- Comment gère-t-on les silences - les réponses qui n'arrivent jamais -, les personnes en présence physique, les personnes en présence médiatique ?
- Existe-t-il une compétence émotionnelle qui nous permette de mieux gérer verbalement nos rapports aux autres ?
- Comment gère-t-on nos addictions à ces nouveaux modes de communication ?
- Existe-t-il une bonne façon de communiquer pour assurer une « popularité médiatique » et pour mieux gérer les situations embarrassantes ?
- Quelles sont les raisons qui poussent les personnes à utiliser ces messageries dans le cadre de leur travail (personnel – professionnel) ?
- En quoi l'outil réseau social facilite-t-il ou contraint-il les interactions ?

MÉTHODOLOGIE DE COLLECTE

Pour répondre à ces questions, nous avons lancé un appel à participation de volontaires qui a duré approximativement six mois. Les données continuent à nous parvenir, malgré le fait que l'appel soit clôturé. Nous demandons à ces volontaires de nous « faire don » de leurs messages en les envoyant à une adresse électronique (pour Facebook, WhatsApp et Viber) et à répondre à deux questionnaires (sociolinguistique et psychologique). Le 1^{er} questionnaire a permis de récolter des données sociodémographiques des participants ainsi que quelques-unes de leurs opinions et pratiques quotidiennes. Le 2^e questionnaire (psychologique) était constitué d'un test de quotient émotionnel élaboré par nos partenaires psychologues et l'entreprise privée en comparant les résultats des participants à ceux de 16 000 personnes représentant un échantillon de la population : ce test permet de dresser une carte des compétences émotionnelles des participants. Parallèlement à l'appel à participation auprès de la population dans le but de collecter des données textuelles issues des réseaux sociaux et messageries électroniques, nous avons également lancé un appel aux professeurs de Communauté française de Belgique ainsi qu'en France afin d'obtenir des productions de langue « normée » des élèves en situation scolaire.

Jusqu'à présent, un millier de locuteurs francophones a répondu à la collecte. Certains à l'une ou l'autre parties décrites ci-dessus, d'autres à toutes. Les discussions Facebook, WhatsApp et Viber que nous avons collectées constituent un corpus d'analyse de 23 millions de mots de discussions entre 11 000 locuteurs. Concernant les données sociodémographiques et psychologiques, 700 personnes

ont répondu aux deux questionnaires. Enfin, 120 élèves ont envoyé des dictées et des préparations de travail en classe.

PREMIÈRES ANALYSES

Même si la collecte pourrait prendre plus d'ampleur afin de travailler sur un échantillon de population plus important, les données collectées jusqu'à présent nous permettent de dresser de premières tendances provenant du questionnaire sociolinguistique ainsi que des conversations Facebook principalement.

Facebook est le réseau social principal pour 70 % de la population interrogée, mais pour les moins de 17 ans, ce réseau mondial (54 %) partage la vedette avec Instagram et Snapchat (34 % ensemble), tandis que pour les 30-35 ans, il remporte 84 % des inscriptions. Au-delà de la fréquence d'utilisation des réseaux sociaux, nous avons également enquêté sur la place psychologique que ces réseaux occupaient : pour seulement 20% des personnes interrogées, les réseaux sociaux jouent une part importante dans leur réputation générale (voire très importante pour 5%), contre 30% pour une part moyenne et 50% pour une part faible.

Concernant les représentations linguistiques, 63 % de la population interrogée se dit « énervée » par les fautes d'orthographe dans les médias numériques et, pourtant, 89 % d'entre ces derniers affirment distinguer clairement faute d'orthographe et variation du code de type « langage sms ». On note cinq raisons principales à cet « énervement » :

1. lecture (18 %) : parce que le message devient difficile à déchiffrer.
2. norme (15 %) : parce qu'il faut connaître et respecter la langue.
3. manque de bonne volonté (14 %) : parce qu'ils pourraient faire attention et ne le font pas.
4. danger (11 %) : parce qu'on va détruire la langue sur le long terme.
5. politesse (9 %) : parce que c'est un manque de respect pour l'interlocuteur.

Au niveau des relations sociales et des états émotionnels, notre questionnaire rapporte que seuls 2 % des personnes interrogées ont souvent souffert de harcèlement sur les réseaux sociaux, 11 % rarement et 87 % jamais. Cependant, les réactions émotionnelles issues des interactions sont plus nuancées. Le graphe suivant résume les principales émotions et leur fréquence chez les personnes à qui nous avons posé la question *Sur les réseaux sociaux, avez-vous déjà ressenti X suite à un post d'une de vos connaissances ?* (X étant le sentiment, parmi les sentiments suivants : embarras, tristesse, oppression, colère ou mal-être qui a ensuite persisté dans la vie réelle).

63 % de la population interrogée se dit « énervée » par les fautes d'orthographe dans les médias numériques

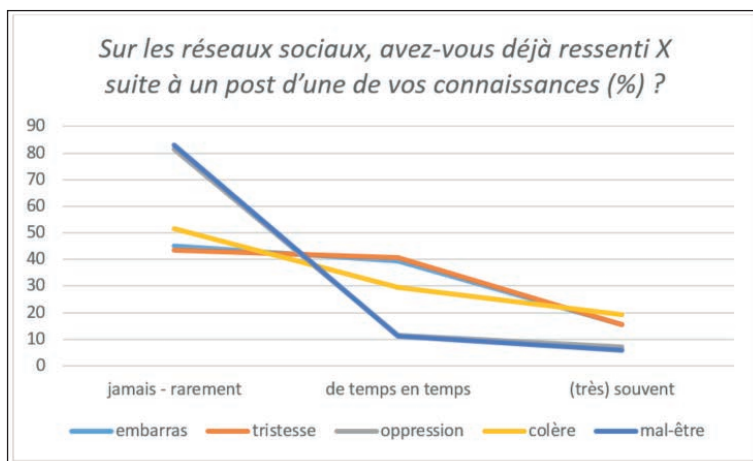


Figure 1. Résultats à la question *Sur les réseaux sociaux, avez-vous déjà ressenti X suite à un post d'une de vos connaissances ?* (réponses possibles : embarras, tristesse, oppression, colère ou mal-être qui a ensuite persisté dans la vie réelle). Projet *Vos Pouces pour la Science*

Au vu des tendances exprimées par les lignes, nous pouvons conclure que l'oppression et le mal-être qui persiste dans la vie réelle sont deux émotions moins fréquentes. Au contraire, colère (19% souvent), tristesse (16% souvent) et embarras (16% souvent) sont les émotions les plus récurrentes suite aux publications de l'entourage de la personne interrogée.

Enfin, en ce qui concerne les analyses linguistiques, nous avons utilisé le logiciel Tropes afin de nous permettre de dresser les premières tendances. Celles-ci sont de trois types : grammaticales, argumentatives et référentielles (thématiques).

(1) Du point de vue des catégories grammaticales¹, les verbes ont une place prépondérante dans les conversations privées de réseaux sociaux. On recense principalement des verbes factifs (40%) et statifs (35%), contre à peine 25% pour les verbes déclaratifs et performatifs. Ce résultat signifie que les locuteurs ont plutôt tendances à raconter des faits (*j'ai commencé mes études*), organiser des événements (*on fera ça demain*) et à parler de leur état (*tu es maso, je vais mieux...*), plutôt qu'à prendre position ou à s'engager. Les connecteurs sont également nombreux, principalement les connecteurs d'addition (33%), d'opposition (25%) et de cause (15%). Nous ne sommes pas étonnée par ce résultat puisque les conjonctions *et*, *ou* et *mais* font les articulations de conversations portant sur des faits et des états. Viennent ensuite les adverbess et les adjectifs. Pour les premiers, il s'agit principalement d'adverbess de négation (24%) et d'intensité (33%) ; pour les seconds, c'est la classe des adjectifs subjectifs (*pratique, perspicace, cool*) qui l'emporte largement (50%) sur celle des adjectifs objectifs (*supplémentaire, final, suivant*). Ce dernier résultat reflète la nature même des conversations sur les réseaux sociaux, aux contenus généralement privés.

(2) D'un point de vue argumentatif, on note une large prise en charge du narrateur (le *je* et le *nous*) qui s'adresse à une personne connue et proche (le *tu* par opposition

¹ Les catégories énoncées ici sont toutes issues du logiciel Tropes.

au *vous* de politesse). Le style des conversations Facebook est plutôt argumentatif « qui discute, qui compare ou qui critique »², plutôt que narratif « qui raconte un récit, à un moment donné, en un certain lieu » (le style de l'œuvre de fiction, par exemple), énonciatif « qui établit un rapport d'influence ou révèle un point de vue » (le style de l'article de presse, par exemple) ou descriptif « qui décrit quelque chose ou quelqu'un ». On aurait pu s'attendre à une affluence de cette dernière catégorie étant donné le type de verbes mentionné précédemment. Ce résultat indique que les conversations sur les réseaux sociaux comprennent bien des verbes factifs et statifs, mais dans l'objectif final de constater, comparer ou critiquer, plutôt que de simplement décrire.

(3) Enfin, d'un point de vue référentiel ou thématique, nous remarquons cinq catégories référentielles récurrentes : la catégorie la plus présente est celle des comportements et sentiments (*triste, amour, pauvre type*), suivie de celle des références temporelles (*Noël, juillet, semaine, vacances*), des personnes et groupes sociaux (prénoms, *gens, clan, famille*), de la communication et des médias (*gsm, mail, logo, site web*) et de l'alimentation et de la vie pratique (*lunch, chocolat, souper, crudités*). Détaillons la première catégorie, « comportements et sentiments », qui est de loin la plus importante, dans l'arborescence suivante.

1 accord et désaccord	4 comportements	5 dérision et ironie
accords	harcèlement	ironie
désaccords	audace	moquerie
réclamation	exagération	
2 agressivité et brutalité	fainéantise	6 jugements et appréciations
agressivité et violence	gentillesse	estimation et diagnostic
coup	honnêteté	appréciation
inhumanité	justification	mérite
méchanceté	malhonnêteté	jugements de valeur
sacrifice	médiance	bête
supplices	nervosité	déguelasse
3 cognition et connaissance	patience et impatience	désagréable
cognition	ruse et hypocrisie	bon
compétence	sociabilité	cool
connaissance	politesse	drôle
état d'esprit	remerciement	génial
mémoire et oubli		intéressant
raisonnement et hypothèse		juste
		parfait
		sérieux
		sympathique
		perception

Figure 2. Thématiques récurrentes dans le cadre référentiel des comportements et sentiments. Projet *Vos Pouces pour la Science*

Nous voyons que, dans l'arborescence, les champs lexicaux sont numérotés et les sous-domaines les plus fréquents leur sont associés. Si l'on prend l'ensemble des sous-domaines indiqués (49), on remarque que pas moins de 22 sous-domaines ont une connotation négative, que ce soit pour critiquer une personne dans son comportement même (*fainéantise, bête*) ou dans sa relation aux autres (*agressivité, harcèlement, médiance, moquerie*). C'est donc la polarité de sentiments la plus récurrente. Peu de sous-domaines sont neutres (7) : cognition, connaissance, état d'esprit, raisonnement et hypothèse, justification, estimation et diagnostic, et perception. 20 sous-domaines sont positifs et se réfèrent à une valeur personnelle (*drôle, sympathique*), à une qualité universelle (*compétence, appréciation*) ou à une qualité interpersonnelle (*gentillesse, sociabilité, politesse*). Le sous-domaine le plus représenté, toutes catégories confondues, est celui des jugements de valeurs.

PERSPECTIVES

² Les définitions de cette partie sont issues du logiciel Tropes.

De nombreuses études peuvent se baser sur les larges corpus et les résultats des questionnaires issus du projet *Vos Pouces pour la Science*. De notre côté, nous avons mis en chantier plusieurs pans de travail de nature diverse que nous allons traiter. Nous allons premièrement approfondir le champ des comportements et sentiments véhiculés dans les réseaux sociaux, et ce, par le truchement de deux études :

- une étude qui vise à comparer le lexique et les tournures employées chez des personnes au quotient émotionnel totalement distincts (en croisant les données des cartes émotionnelles des participants avec leurs conversations Facebook) ;
- une étude qui vise à saisir la nature et l'intensité des conversations à polarité négative, dont principalement celles empreintes de violence verbale (insultes, harcèlement psychologique et sexuel) ; ceci se fera par une recherche automatique des lexiques polarisés et une analyse qualitative des candidats en contexte.

Ensuite, nous travaillons sur la prégnance des langues étrangères dans une situation de langue principale française : nous voulons, pour un échantillon des 22 millions de mots collectés, étudier la place et les fonctions des mots étrangers utilisés en français : rôle emphatique, ludique ou culturel ? L'étude de cette alternance de langue permettra également de comparer les résultats obtenus dans les réseaux sociaux avec ceux obtenus précédemment pour les sms uniquement (Cougnon, 2011).

Enfin, nous voulons étudier la variation orthographique des 120 élèves âgés entre 13 et 17 ans qui ont participé au projet, en comparant leur production écrite dans une situation formelle (dictée et préparations de travail en classe) avec celle issue des réseaux sociaux. L'objectif ici est de quantifier le taux d'erreurs orthographiques par rapport aux jeux sur la langue et à la langue normée. Ce quatrième pan (déjà entamé dans Maskens *et al.*, 2015) a pour but de répondre aux questionnements des parents, professeurs et journalistes dans une dynamique scientifique de service à la société.

RÉFÉRENCES

Louise-Amélie Cougnon (2011). « Tu te prends pour the king of the world? ». Language contact in text messaging context ». In Cornelius Hasselblatt, Peter Houtzagers and Remco van Pareren (eds). *Language Contact in Times of Globalization. Studies in Slavic and General Linguistic*, 38, Rodopi, Amsterdam/New York, pp. 45-59.

Lénaïs Maskens, Louise-Amélie Cougnon, Sophie Roekhaut and Cédric Fairon (2015). « Nouveaux médias et orthographe. Incompétence ou pluri-compétence ? » In *Discours. Revue de linguistique, psycholinguistique et informatique* 16 [online]

Louise-Amélie Cougnon est linguiste, chargée de recherche au FNRS. Après avoir obtenu une licence en linguistique à l'UCL, Louise-Amélie a réalisé une thèse de doctorat portant sur la communication par sms. Depuis, sa recherche se centre sur la linguistique du numérique et sur ses liens avec d'autres disciplines : le discours politique sur Twitter, les compétences émotionnelles sur Facebook, les représentations des locuteurs par rapport aux jeux sur la langue dans les sms... autant de sujets pluridisciplinaires au centre de ses recherches.

