

Algospeak

Description

Néologisme pour désigner le langage créé par les utilisateurs des médias sociaux afin d'échapper aux filtres de la modération des contenus pratiquée sur ces plateformes ([voir La rem n°59, p.90](#)), l'*algospeak* se décline en initiales, mots aux lettres inversées ou modifiées, acronymes, euphémismes, métaphores, allographes (lettres à lire les unes après les autres). Ces figures de style travestissent les mots et les expressions dont l'emploi entraîne la suppression ou la mise en invisibilité (*shadowban*) des messages postés notamment sur Facebook (2,9 milliards d'utilisateurs), YouTube (2,5 milliards), Instagram (1,5 milliard) ou TikTok (1 milliard).

Dès le début de l'ère numérique, les diverses technologies ont contribué à forger de nouveaux modes d'expression, éléments fédérateurs des usages internet. Dans les années 1980, les programmeurs en informatique parlaient le *leet speak* (*elite speak* en anglais). Ils avaient inventé ce langage afin de communiquer entre initiés sur les premiers services en ligne d'échange de fichiers appelés BBS (*Bulletin Board System*), ancêtres des forums qui les remplaceront avec l'avènement de l'internet. Écriture construite sur la ressemblance des caractères – le « S » devient « 5 » ; le « K » s'écrit « |< » –, le *leet speak* est encore utilisé aujourd'hui pour diffuser des courriels indésirables (*spam*). La généralisation des conversations en ligne sur les forums et les chats a favorisé l'emploi des abréviations finalement intégrées au langage courant en ligne ou hors ligne, comme LOL (*laughing out loud*/ mort de rire), ASAP (*as soon as possible*/ dès que possible), FYI (*for your information*/ à titre d'information) ou encore IRL (*in real life*/ dans la vie réelle). Au début des années 1990, le langage SMS est né d'une contrainte technique : un nombre limité de caractères qui a lancé les raccourcis phonétiques comme *jamè, grav, eske, koi 2 9 ?* (jamais, grave, est-ce que, quoi de neuf ?). Sur le réseau social Twitter, avec la contrainte des 280 signes par tweet, on discute aujourd'hui par *hashtag* (mot clé) et en *thread* (fil de messages).

L'*algospeak*, quant à lui, est né pour duper les machines. Plus de la moitié de la population mondiale utilise les réseaux sociaux sur lesquels des machines sont programmées pour lutter contre les conversations infâmes en tout genre – propos extrémistes, contenus nuisibles, applications malveillantes – et pour combattre autant que possible le puits sans fond de la désinformation. Mais cette traque automatisée n'est pas sans failles. Avec des algorithmes qui opèrent selon une logique systématique manquant forcément de subtilité, la modération sur les plateformes internet finit par évincer des sujets de société majeurs et par rendre inaudible la parole des minorités ou des communautés marginalisées. Parler en *algospeak* devient une parade à la stigmatisation. Les personnes de couleur et les personnes transgenres se protègent avec ce lexique plus discret pour dire l'oppression dont elles sont victimes, évitant d'employer les mots « raciste » ou « blanc ». Pour les créateurs de contenus, appelés « influenceurs », l'adoption de ce langage codé est un moyen de prévenir le risque de suppression de leurs posts en raison d'expressions ou de

sujets catalogués comme « inappropriés » par les algorithmes, et de préserver ainsi leurs revenus publicitaires. La création de ce subterfuge langagier illustre bien les limites de la modération automatique des contenus à grande échelle.

Le 8 avril 2022, *The Washington Post* titre, sous la plume de Taylor Lorenz : « *L'algospeak d'internet change notre langage en temps réel.* » De nouveaux vocables se multiplient en ligne, annonce le quotidien américain. Le mot « pandémie », sujet majeur de l'actualité récente, se trouve particulièrement scruté par les algorithmes dans la lutte contre la désinformation. Le mot est alors transcrit en *algospeak* par les mots clés « panda express » ou « panini » sur les médias sociaux. En riposte à la décision de YouTube de démonétiser les vidéos mentionnant le coronavirus, les créateurs de *Game Grumps*, web séries de jeux vidéo, ont choisi, quant à eux, la formule « Backstreet Boys Reunion Tour » pour parler de la pandémie, faisant référence à « *l'événement mondial affectant la vie des gens* » que représente la tournée du célèbre boys band. Autre exemple, les groupes antivaccins sont appelés « *dance party* » ou « *dinner party* », tandis que les personnes vaccinées sont des « nageurs ».

« *L'algospeak devient de plus en plus courant sur internet, car les gens cherchent à contourner les filtres de modération de contenus sur les plateformes de médias sociaux* » explique encore la journaliste Taylor Lorenz. Sur les médias sociaux, on ne parle pas librement de sexe, d'avortement, de violence, de guerre, de viol, d'addiction. De nombreux sujets sont inscrits sur une liste noire surveillée par les machines. Sur TikTok, les « comptables » désignent les travailleurs du sexe, avec l'emoji représentant un épi de maïs (*corn*) pour *porn*. « *Cornucopia* » remplace le mot « homophobie » et « *leg booty* » signifie LGBTQ. Dans les communautés féministes, toujours sur TikTok, l'*algospeak* traduit le mot viol par « vi01 », violence par « v10lence », ou encore *lesbian* par « le\$bian » [prononcé « le dollar bean »]. La santé des femmes est également un thème qu'il vaut mieux aborder en langage codé en écrivant « s€in » pour sein, « r€gles » pour règles ou éducation « seggsuelle ». « *Cela donne l'impression que vous n'êtes pas professionnel d'avoir ces mots orthographiés bizarrement dans vos légendes, regrette une créatrice de contenus. Surtout pour un contenu qui est censé être sérieux et à caractère médical.* »

En *algospeak*, les euphémismes sont fréquents – « *unalive* » pour *dead*, « *opposite of love* » pour *hate* ou « devenir non vivant » pour suicide, tout comme les emojis et les symboles couramment employés pour remplacer une ou deux lettres d'un mot (« *depre\$\$ion* » pour *dépression*), juste assez pour passer sous le radar de la censure automatique. Dans les discussions sur YouTube et TikTok à propos de l'invasion du pays par la Russie, l'Ukraine est citée à l'aide d'un emoji représentant un tournesol (en référence à la pénurie d'huile). Comme l'indique Thibault Prévost sur le site Arrêt sur images, même le youtubeur Hugo Travers utilise l'*algospeak* sur sa chaîne d'actualité HugoDécrypte pour son 1,5 million d'abonnés, parsemant les légendes de ses vidéos de symboles afin d'éviter les mots repérables par les algorithmes de modération.

C'est sur TikTok, plateforme vidéo d'envergure mondiale appartenant au groupe chinois ByteDance, que la modération des contenus est la plus stricte. La visibilité des contenus dépend exclusivement de l'algorithme qui remplit la rubrique « *For You* », sélection des contenus adaptée à chaque utilisateur. Aussi sévère

qu'opaque, l'algorithme refuse d'entendre parler, pêle-mêle, d'homophobie, de LGBTQ+, de règles, de viol, de dépression, selon le *Washington Post*. Faire une recherche sur TikTok concernant un trouble alimentaire comme l'anorexie renvoie à la rubrique « Parles-en à un professionnel ». Dans la liste des règles communautaires à respecter, la plateforme chinoise rappelle que son système de recommandation personnalisé vise à « *contrer certains des problèmes auxquels tous les services de recommandation peuvent être confrontés* » et elle fournit des conseils « *pour personnaliser votre expérience de découverte* ».

De leur côté, les créateurs de contenus, par l'intermédiaire de la TOCA – The Online Creators' Association –, attendent de TikTok un mécanisme de signalement et de recours plus transparent et plus réactif qui impliquerait des êtres humains. L'objectif est de remédier aux effets pervers du systématisme des signalements de masse. À partir des informations recueillies auprès des utilisateurs de la plateforme, la TOCA s'interroge notamment sur le déploiement erratique des nouvelles fonctionnalités, qui peuvent disparaître sans prévenir ou être proposées à un nombre limité de créateurs, sans rapport avec la localisation ou le type d'appareil utilisé.

Les créateurs de contenus partagent des listes de plusieurs centaines de mots bannis par les systèmes de modération, raconte le *Washington Post*. « *Mémorial dédié aux mèmes qu'Instagram juge inadaptés à sa plateforme* », le site Zuck Got Me For est une illustration de cette dissidence qui s'organise. Son administratrice @NeoliberalHell2 s'en explique dans un manifeste en ligne : « *La liberté créative est l'un des seuls aspects positifs de cet enfer en ligne dans lequel nous vivons tous. [...] Avec des règles de plus en plus strictes et des algorithmes plus rigoureux, les créateurs, les memers et les satiristes sont plus que jamais censurés. [...] Ces mesures plus sévères créent des menaces majeures pour la liberté d'expression et sont préjudiciables aux créateurs, dont beaucoup paient leurs factures en monétisant les applications.* » Elle raconte également l'impossibilité pour les créateurs de contenus d'obtenir le soutien du géant internet. En 2016, 175 créateurs de contenus comptant plus de 20 millions de *followers* et de *likes* sur Facebook avaient tenté, sans succès, un « Zuxit », en boycottant l'application pour obtenir un processus transparent de signalement et de suppression sur le plus grand réseau social du monde. En 2019, le groupe @UnionizedMemes revendiquait, à son tour, « *un processus de recours plus ouvert et transparent pour les interdictions de compte* » et, notamment, « *une liaison dédiée à la communauté des mèmes* ». L'administratrice, qui délivre des informations et des conseils sur le dysfonctionnement de la modération, déplore « *l'incompétence générale de ces algorithmes qui rend la situation si préjudiciable aux créateurs en ligne* », réclamant un service apte à trancher avec plus de nuances qu'une machine.

Enseignante à la faculté de droit de l'UCLA (University of California, Los Angeles) et spécialiste des liens entre technologie et discrimination raciale, Ángel Díaz fait cette mise au point dans son entretien avec le *Washington Post* : « *La réalité est que les entreprises technologiques utilisent des outils automatisés pour modérer le contenu depuis très longtemps et, bien qu'ils soient présentés comme un apprentissage automatique sophistiqué, il s'agit souvent d'une simple liste de mots qu'ils considèrent comme problématiques.* »

À la tête de Fight for the Future, groupe de défense des droits numériques, Evan Greer, quant à elle, insiste

sur l'importance des dommages collatéraux de la modération automatique, sachant que ceux qui agissent avec malveillance trouveront toujours la parade. « *Tenter de réguler le discours humain à l'échelle de milliards de personnes dans des dizaines de langues différentes et essayer de faire face à des choses telles que l'humour, le sarcasme, le contexte local et l'argot ne peut pas se réduire au simple classement de certains mots* », déclare-t-elle au *Washington Post*. D'autant que les progrès de l'apprentissage automatique (*machine learning*) rendent la tâche encore plus difficile : quand, victimes de leur notoriété, les termes de l'*algospeak* deviennent un argot courant, ils n'échappent plus à leur tour aux filtres des machines.

À travers le phénomène *algospeak* apparaissent les limites d'un système de modération devenu insensé car exagérément censeur. Leur gigantisme, par la masse d'utilisateurs, impose aux médias sociaux de recourir à l'intelligence artificielle pour bannir les contenus nuisibles ou illicites, au risque d'araser toute nuance, toute différence, toute créativité et d'ébrécher la liberté de parole. « *Si l'on opte pour la voie humaine, un réseau social n'aura plus besoin de s'appuyer sur le bannissement comme premier (et seul) mécanisme de normalisation* » selon Clive Thompson, journaliste au *New York Times Magazine* et à *Wired*, qui préconise dans un article publié sur OneZero (Medium), cité par Arrêt sur images, « *d'appliquer un jugement humain au comportement des utilisateurs* ». Si leur jugement est faillible, les hommes n'en sont pas moins responsables et ils peuvent rendre des comptes. Les décisions algorithmiques sont « *désormais si complexes que les ingénieurs eux-mêmes ne peuvent pas toujours expliquer pourquoi un contenu est déclassé* », explique Clive Thompson. Les employés des réseaux sociaux qui effectueront ce travail de modération devront être correctement payés ; pas question d'« *externaliser ce travail à des fermes de modération de contenu tierces aux conditions de travail épouvantables* ». Pour infléchir la logique de la modération des contenus en remplaçant les robots par des humains convenablement rémunérés, encore faudrait-il que leurs utilisateurs ne se comptent pas par milliards. Et Clive Thompson de conclure : « *Il s'agit, en réalité, d'un argument de plus en faveur d'un examen antitrust des principaux réseaux sociaux. Si l'un des problèmes majeurs liés à leur incapacité à modérer leur contenu avec un jugement à l'échelle humaine est qu'ils sont tout simplement trop grands, peut-être que c'est la taille qui pose problème.* »

Sources :

- The Online Creators' Association, officialtoca.com.
- « Internet “algospeak” is changing our language in real time, from “nip nops” to “le dollar bean”, Taylor Lorenz, washingtonpost.com, April 8, 2022.
- *Leet speak*, Wikipédia, <https://fr.wikipedia.org/>, dernière modification de cette page le 16 avril 2022.
- « Of “Algospeak” and the Crudeness of Automated Moderation », Clive Thompson, OneZero, onezero.medium.com, April 15, 2022.
- « Algospeak : les influenceurs, bilingues en censure automatisée », Thibault Prévost, arretsimages.net, 8 mai 2022.

Categorie

1. A retenir

date créée
juillet 2022

Auteur
francoise