# Les vingt premières années du capitalisme linguistique : Enjeux globaux de la médiation algorithmique des langues

Frédéric Kaplan, DHLAB, EPFL. / frederic.kaplan@epfl.ch

La médiation des flux linguistiques par une poignée d'acteurs mondiaux a permis la constitution de modèles de la langue dont les performances sont aujourd'hui sans précédents. Ce texte revient sur les thèses principales qui expliquent des dynamiques économiques et technologique à l'origine de ce phénomène, prolongeant les réflexions préalablement entamées [1] [2] (thèses 1 et 2), pour ensuite développer la question du développement et de l'usage des nouveaux modèles basé sur le capital linguistique accumulé (thèses 3 et 4) et les conséquences sur les nouvelles formes de médiation algorithmiques qui en découlent (thèse 5 et 6). L'enjeu est de comprendre le basculement qui s'opère dans l'économie de l'expression quand les algorithmes produisent des écrits de performativité supérieure à la langue courante et sont donc naturellement utilisés pour produire du texte par l'intermédiaire de prothèses. Nous assistons à ce que l'on pourrait qualifier de "second régime du capitalisme linguistique" allant au-delà de la vente aux enchères des mots pour proposer additionnellement la vente de services linguistiques se substituant à l'écriture même. Il devient alors possible, à l'image des services de traduction automatique, mais cette fois-ci pour sa propre langue, de s'exprimer avec plus d'efficacité et de produire de l'écrit adapté aux diverses exigences des situations de la vie professionnelle ou privée. Nous discutons comment la généralisation de ces prothèses linguistique pourrait conduire à une atrophie de la capacité expressive, pour l'écriture dans un premier temps et, peut-être, à terme pour l'oralité.

## Thèse 1 : Le capitalisme linguistique invente un régime économique caractérisé par **la monétisation des langues au niveau mondial**.

En 1998, Sergueï Brin et Larry Page développent une **valeur d'usage** pour Google en inventant l'algorithme *PageRank* qui classe les pages webs en fonction des pages qui les citent et de leur importance. Ils présentent le principe de cet algorithme inspiré par la méthode d'évaluation de l'impact des publications académiques à une conférence en Australie puis publient le premier article sur le sujet dans la revue issue des actes de la

[1] Kaplan, F. (2011, November). Quand les mots valent de l'or. Le Monde Diplomatique, 28.

[2] Kaplan, F. (2014). Linguistic Capitalism and Algorithmic Mediation. Representations, 127(1), 57–63. https://doi.org/10.1525/rep.2014.127.1.57

conférence[3]. Pour mesurer la pertinence d'une page web à une requête, *PageRank* résiste mieux que les autres mesures basées sur simplement la présence des mots clés. Il permet à Google de rapidement s'imposer comme le meilleur moteur de recherche de cette époque.

Mais durant les premières années de l'entreprise, Google doit encore trouver une **valeur d'échange** pour son service. En mars 2000, la "Bulle Internet" éclate. Beaucoup de « start-ups » ne proposant que des services attrayants, mais pas de sources stables de revenus, ne résistent pas. C'est le cas en particulier des entreprises qui misaient seulement sur l'affichage de la publicité dans leurs espaces "virtuels" sur un modèle simplement transposé de celui du monde réel.

Le modèle *AdWords* de Google lancé le 23 mars 2000 avec seulement 350 partenaires (ils sont plusieurs millions aujourd'hui), est différent. Brin et Page ont réalisé qu'ils interceptaient un **flux linguistique** grâce au nombre croissant d'utilisateurs qui cherchaient au travers de services de Google. Google développe alors un algorithme qui organise automatiquement la spéculation autour des mots recherchés. Le principe de l'algorithme qui gère les enchères est le suivant. Une entreprise choisit un mot ou une expression, par exemple « vacances » et le prix maximum qu'elle serait prête à payer en cas de clic. Google propose une estimation E du montant de l'enchère à proposer pour avoir de bonnes chances d'être dans la première page des résultats proposés. Google donne aussi un score à la publicité sur une échelle de un à dix. Ce score dépend essentiellement de la pertinence du texte de la publicité par rapport à la requête de l'utilisateur, de la qualité de la page vers laquelle la publicité pointe (qualité de son contenu et rapidité de chargement) et du nombre de clics moyen sur la publicité. En gros, ce score mesure à quel point la publicité « fonctionne ». Le rang R de l'annonce est simplement l'enchère E multipliée par le score S. Une publicité ayant un bon score peut ainsi compenser une enchère plus faible et arriver devant.

Ce jeu d'enchères est recalculé pour chaque requête de chaque utilisateur — des millions de fois par seconde. Le marché linguistique ainsi créé par Google est très vite global et multilingue. La « Bourse des mots » qui lui est associée donne une indication relativement juste des grands mouvements sémantiques mondiaux.

Le marché linguistique existait déjà avant Google. Il était possible depuis longtemps d'acheter certains mots, comme des noms de marques. L'algorithme de Google a élargi et en partie libéralisé ce marché. Tous les mots pouvant maintenant donner lieu à des enchères, selon un mécanisme transparent.

Vingt ans plus tard, c'est encore ce mécanisme qui assure les revenus et la croissance de Google, maintenant filiale d'un groupe nommé (de façon transparente) Alphabet. Au sein de groupe, les revenus de Google représentent plus de 99% du chiffre d'affaires. Pour avoir un ordre de grandeur, Google réalisait en 2020, 180 milliards de dollars de chiffre d'affaires par an, soit 491 Million de dollars par jour, plus de 20 millions par heure. Ce chiffre est obtenu essentiellement en organisant la vente des mots à l'échelle planétaire. L'histoire passée du

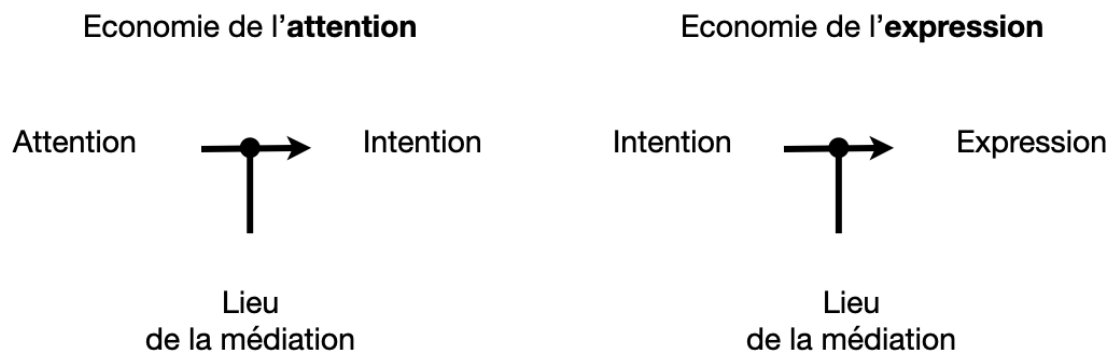
[3] Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems, 30(1), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

développement et des innovations de Google doit donc se relire selon ce prisme, ses choix futurs également.

Le capitalisme linguistique est un territoire inédit du capitalisme qui a été découvert et immédiatement exploité de manière monopolistique par Google. Cela ne veut pour autant pas dire que Google gardera à jamais sa position dominante. Le capitalisme linguistique est caractérisé par plusieurs régimes, dont la vente aux enchères de mots ne constitue qu'une première dimension. Par ailleurs le capitalisme linguistique modifie lui-même les objets qu'il exploite, c'est-à-dire les langues créant ainsi différentes formes d'instabilités dont les thèses suivantes explorent quelques aspects.  À ce jour cependant, Google nous fournit le meilleur exemple d'une entreprise ayant réussi à l'échelle planétaire en maîtrisant les règles de ce nouveau territoire économique.

## Thèse 2 : Le capitalisme linguistique, une **économie de l'expression**, conduit à l'**écriture régulée**.

Le capitalisme linguistique n'est pas une économie de l'attention, mais une économie de l'expression. L'objectif n'est tant de capter les regards, mais d'être un **médiateur** de l'écrit et de l'oralité. Dans une économie de l'attention, le lieu de la médiation est entre l'attention et l'intention. Dans une économie de l'expression, le lieu de la médiation est entre l'intention et l'expression. Ceci revient à dire que dans une économie de l'expression, il faut développer des relations linguistiques intimes et durables avec un grand nombre d'utilisateurs, pour pouvoir capturer, modéliser et infléchir la langue.




*Économie de l'attention et économie de l'expression*

Dès les premières années du déploiement des services autour de son moteur de recherche, quand Google corrigeait à la volée un mot que vous aviez mal orthographié, il transformait

un matériau sans valeur en une ressource économique potentiellement rentable. Quand Google prolongeait une phrase que vous avez commencée à taper, il ramenait votre expression dans le domaine de la langue qu'il pouvait exploiter. Le mot "Picassso" écrit avec trois "sss" n'a aucune valeur marchande. Corrigé en "Picasso", le mot est immédiatement associé à des enchères importantes ayant pour objectif de proposer des visites de musée ou l'achat d'une voiture. D'une manière plus générale, ces premiers dispositifs de médiations permettaient déjà de **régulariser la langue**, de la rendre prédictible par des modèles, pour maximiser son potentiel économique.

Entre les années 2010 et 2020, l'autocomplétion s'est rapidement généralisée sur presque toutes les interfaces de saisie. Aujourd'hui dans bien des situations, comme avec la fonctionnalité de "rédaction intelligente" proposé par le service *Gmail*, il ne s'agit plus de taper, mais de **diriger** son texte dans une direction ou une autre, comme lorsque l'on conduit un véhicule. Nous choisissons un chemin parmi l'arbre des expressions possibles. Nous pouvons parler d'**écriture régulée**.

Cette transformation de la dynamique d'écriture est bien plus profonde que celles qui ont caractérisé le passage de l'écriture manuelle à la machine à écrire puis au traitement de texte. Marshall McLuhan notait qu'en 1882, la publicité proclamait que la machine à écrire était d'un précieux secours dans l'apprentissage de la lecture, de l'écriture, de l'orthographe et de la ponctuation. Elle forçait l'explicitation de la nature écrite de la langue, notamment par la procédure de la dictée. Elle ralentissait le passage de l'oral à l'écrit, forçant l'explication de son articulation [4]. C'est cette même année que Nietzsche acheta une des premières machines à écrire alors qu'il éprouvait depuis plusieurs mois de grandes difficultés pour se mettre au travail, traversant une période de grave dépression. Avec un peu d'entraînement, Nietzsche réussit à taper les yeux fermés et recommença à écrire avec plaisir. Pourtant, il constata une transformation subtile de son style d'écriture. Ses phrases étaient devenues plus serrées, plus télégraphiques comme si le « fer » de la machine durcissait son discours. Son écriture médiée par la machine mécanique avait changée[5].

Près d'un siècle plus tard, le traitement de texte en donnant notamment la possibilité de corriger, de revenir en arrière a transformé le processus d'écriture en une séquence d'actions réversibles. Chaque texte est devenu une histoire, toujours possiblement en devenir, parfois même une histoire collective, un processus en convergence lente, un chantier perpétuel, à l'image des articles de Wikipedia. Le traitement de texte a aussi ouvert une fenêtre d'observation sur le processus d'écriture lui-même, donnant pour la première fois une possibilité de voir des dynamiques de l'écrit, considérant le texte en composition comme un flux et ouvrant implicitement à la prochaine transformation du procédé d'écriture[6].

---

[4] Marshall McLuhan (1964) Pour comprendre les médias, Point: Seuil, p.297

[5] Cet épisode est cité par Carr, N. (2010). The Shallows: What the Internet Is Doing to Our Brains, Atlantic Press, lui-même reprenant le travail de Kittler, F. A., (1999). Gramophone, Film, Typewriter, Stanford University Press.

[6] Doquet-Lacoste, C. (2009). L'écriture sur traitement de texte: Quelques spécificités de son analyse en temps réel. Modèles linguistiques (59), 133–151. https://doi.org/10.4000/ml.341

Depuis 2010, avec la génération de l'écriture régulée, nous assistons à un raccourcissement du circuit de la production de l'écrit, une accélération de la performativité de l'écriture par une optimisation du processus qui part de l'intention à l'expression. L'histoire, la séquence transparente d'action qui caractérisait le traitement de texte a disparu, caché dans le fonctionnement invisible de l'algorithme. L'explication structurelle et orthographique imposée par la machine à écrire mécanique est oubliée. Une autre forme d'expressivité commence.

De la même manière qu'il y a le regard non encore fixé, il y a l'énoncé non encore formé. Comme dans l'économie de l'attention, l'économie de l'expression court-circuite un processus traditionnellement non médié avec deux objectifs : intercepter un flux et le rediriger selon un critère de pertinence apparente. Les interfaces de médiation permettent une expression plus efficace, sans erreurs orthographiques, dans une langue formellement plus correcte et régulière. La médiation permet des recherches plus efficaces, des courriers plus rapidement écrits, des textes sans erreurs. Elle ne demande plus de virtuosité. Au contraire elle démocratise **la belle écriture**.

Au fur et à mesure que progressent les techniques d'autocomplétion la frontière de la médiation s'éloigne de la correction du texte pour se rapprocher de l'intention, parfois encore indécise de la phrase à produire. Il suffit d'amorcer la dynamique d'écriture pour que des chemins textuels explicitent se matérialisent comme autant de possibles réifications de la proposition initiée. La nature de cette réification est le fruit d'une stratégie algorithmique spécifique invisible pour l'utilisateur qui opère dans un contexte technique et économique particulier. Il convient donc d'ouvrir la boite noire.

## Thèse 3 : Dans le capitalisme linguistique, les modèles linguistiques dérivés de l'interception des flux linguistiques constituent le cœur du capital.

Les acteurs du capitalisme linguistique tendent à engranger des données linguistiques toujours plus nombreuses. Ils peuvent sur cette base produire de vastes modèles de la langue prédictible. Ce capital se transforme en produits et services de médiation linguistique qui à leur tour génèrent plus de capital linguistique. Parallèlement, par l'organisation de la vente des mots ou par la vente de services linguistiques, le capital linguistique en croissance est convertible en capital économique. C'est ce capital économique et les revenus qu'il génère qui permet de soutenir les importants efforts pour extraire des flux linguistiques les modèles linguistiques sous-jacents.

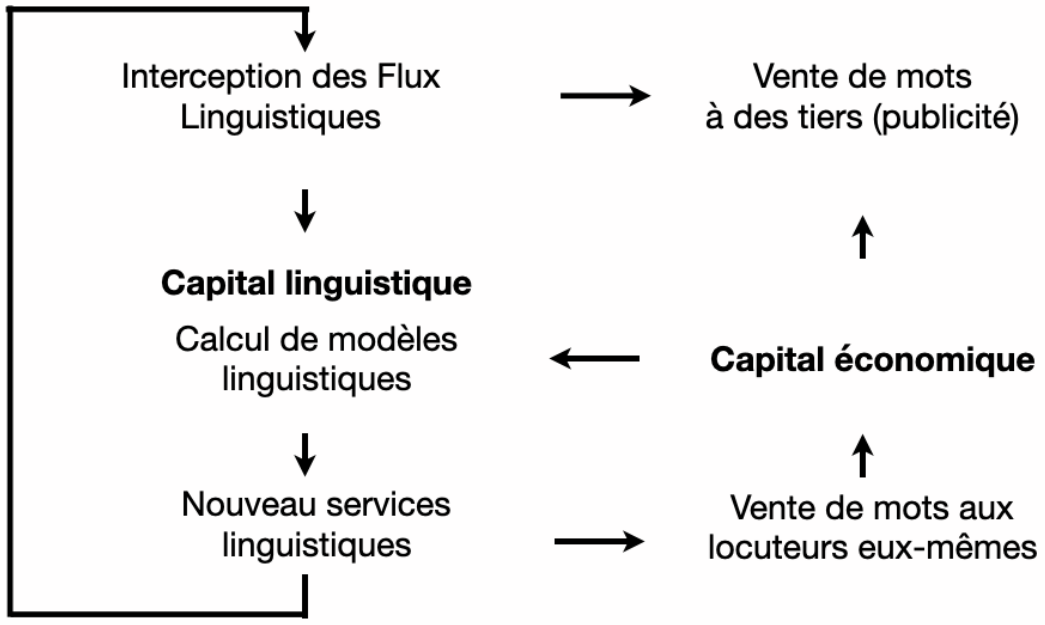



*Schéma général des dynamiques de rétroaction du capitalisme linguistique.*

Pendant des années, la génération de nouveaux textes par des algorithmes était limitée à des écrits cours et formatés pour des contextes très précis. Il était par exemple, possible de programmer des algorithmes pour écrire des articles commentant des résultats sportifs.

En 2017, une équipe de Google développe un premier modèle capable d'apprendre les structures linguistiques en utilisant les mécanismes dits *attentionnels* à large échelle[7]. L'article intitulé *Attention is all you need* est cité plus de 20 000 fois en deux ans et par suite, le premier auteur de cet article, Ashish Vaswani, senior scientist chez Google, devient, presque du jour au lendemain, une des personnes les plus cités de l'histoire de l'informatique.

L'article introduit une architecture appelée *Transformer* qui surpasse, pour certaines taches, les architectures à réseaux de neurones récurrents qui pendant des années constituaient la référence dans l'apprentissage du langage naturel. Pour traduire un texte dans un autre, il faut encoder la première phrase dans une représentation intermédiaire et décoder ensuite sur cette base la représentation pour atteindre la phrase cible. Un *Transformer* met en chaîne une "pile" d'encodeurs puis une "pile" de décodeurs. Chaque encodeur, brique de base du système, est identique,  composé d'un mécanisme dit d'"auto-attention", suivi d'un réseau de neurone qui effectue l'encodage. Les décodeurs utilisent l'attention des encodeurs pour à leur tour décoder les représentations produites vers la phrase cible.

[7] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems, 6000–6010.

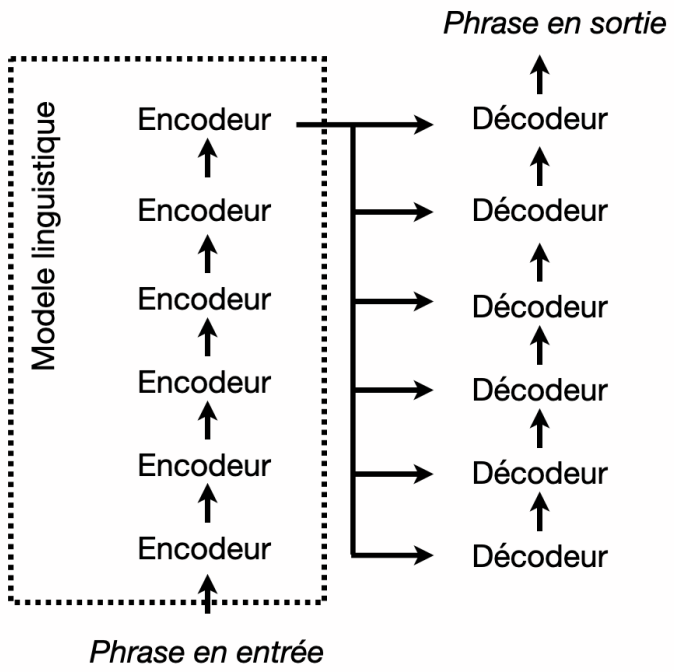



*Schéma simplifié d'un Transformer composé d'une pile d'encodeurs et une pile de décodeurs. Les piles d'encodeurs peuvent servir à construire un modèle linguistique qui peut ensuite être utilisé pour une grande variété de taches.*

Ce mécanisme "attentionnel" est la clé de cette nouvelle architecture. Il donne à chaque traduction successive d'une couche à l'autre un contexte qui module le travail des encodeurs et décodeurs. Les couches attentionnelles hiérarchiques mêlent aussi bien les dimensions syntaxiques que les dimensions sémantiques et peuvent ignorer la forme de toute une partie de la phrase pour ne se concentrer que sur les aspects morphosyntaxiques de certains mots.

Pour entraîner un *Transformer*, il faut lui fournir un grand nombre de paires de phrases (entrée, sortie). À chaque étape de son apprentissage, le *Transformer* fait des prédictions pour les phrases en sortie qui doivent correspondre aux phrases d'entrée. En mesurant la différence entre la phrase de sortie attendue et la phrase prédite, une erreur est calculée et un mécanisme d'apprentissage permet de modifier les poids des réseaux de neurones pour progressivement adapter le comportement de chaque couche d'encodage, chaque couche de décodage et tous les mécanismes d'attentions associés. Si l'on possède beaucoup de données d'entraînement, il devient possible d'obtenir d'excellentes performances.

En 2018, un nouvel article des chercheurs de Google montre que la machine peut produire un modèle de langage et utiliser de l'auto-apprentissage pour améliorer ses performances[8]. En quelques mois, l'article en prépublication sur arXiv est déjà cité plus de 600 fois, avant même d'être accepté. Il est aujourd'hui, comme l'article de 2017, cité plusieurs dizaines de milliers de fois. Le modèle *BERT* (Bidirectional Encoder Representations from Transformers)

[8] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186. https://doi.org/10.18653/v1/N19-1423

reprend le mécanisme d'attention du *Transformer*, qui apprend les relations contextuelles entre les mots d'un texte, mais se concentre sur la partie encodage, indépendamment de toute tache de traduction ou transformation spécifique[9].

La grande révolution du modèle *BERT* est qu'il apprend directement **à partir de flux linguistique** de manière non supervisée. Ainsi, il intériorise des modèles statistiques d'utilisation de la langue potentiellement beaucoup plus large que les modèles précédents. Une fois entraîné, il peut être utilisé pour une très grande variété de taches, en n'ajoutant qu'une petite couche au modèle de base. Il est possible d'effectuer des tâches de classification juste après l'encodage, de reconnaitre des entités nommées (personne, organisation, date, etc.), d'apprendre au système à répondre à des questions, de résumer ou paraphraser un texte, etc. Il est aussi évidemment possible d'utiliser le modèle linguistique ainsi créé pour établir des stratégies plus efficaces dans le cadre du modèle commercial central de Google, la vente des mots. Un modèle comme *BERT* est un véritable capital qui permet de produire ou d'optimiser une très grande variété de services linguistiques.

Les chercheurs de Google sont certes brillants, mais ils ont aussi accès à un capital linguistique sans précédent. Ceci leur permet de rester en tête de toutes les compétitions mondiales dans ce domaine comme la compétition *GLUE* (General Linguistique Undestanding and Evaluation), l'équivalent des Jeux Olympiques pour les modèles linguistiques[10]. Ces compétitions testent les limites des performances des modèles en leur proposant entre autres des tâches de compréhension difficile comme les schémas de Terry Winograd du type "Cette statue ne peut pas rentre dans la boite, car elle est trop petite" ("elle" fait référence à la boite) "Cette statue ne peut pas rentrer dans la boite, car elle est trop grande" ("elle" fait référence à la statue).

Étant donné l'importance des enjeux de ce type de modèles pour produire de nouveaux services, d'autres entreprises sont aussi entrées dans la compétition linguistique. OpenAI est une entreprise à « but lucratif plafonné » dont le président est Elon Musk. Elle est soutenue par ses fondateurs sur un budget global d'un milliard de dollars. *GPT-3* produit par OpenAI est un modèle linguistique basé sur 175 milliards de paramètres[11]. OpenAI ne disposant pas de son propre capital linguistique, *GPT-3* a été entraîné sur des ressources publiquement disponibles ou facilement accessibles, par exemple sur le "CommonCrawl"[12], une archive publique du web contenant des petabytes de données récoltées sur plus de 8

---

[9] Pour apprendre, BERT utilise deux stratégies. Dans la première stratégie qui travaille au niveau des relations des mots dans la phrase, 15% des mots de chaque séquence sont remplacés par un "masque". Le modèle tente ensuite de prédire la valeur originale des mots masqués, sur la base du contexte fourni par les autres mots, non masqués, de la séquence. Dans le second modèle d'entraînement, le modèle reçoit des paires de phrases en entrée et apprend à prédire si la deuxième phrase de la paire est ou non la phrase suivante dans le document original. Dans les modèles d'entraînement utilisés, c'est le cas dans 50% des cas. L'enjeu pour le modèle est donc d'apprendre à modéliser les liens entre phrases du discours.

[10] https://gluebenchmark.com/

[11] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

[12] https://commoncrawl.org/

ans de navigations web, ou sur les contenus de l'encyclopédie en ligne Wikipédia. *GPT-3* n'est entraîné que sur l'Anglais (contrairement à *BERT* de Google qui traite plus de centaine de langues) et d'une manière générale a effectué son apprentissage sur des flux textuels massifs moins contrôlés. Alors que le modèle *BERT* était basé sur 3,3 milliards de mots dans des corpus choisis par les chercheurs de Google, *GPT-3* utilise lui près d'un demi-trillion de mots, soit la moitié d'un milliard de milliards de mots. Malgré le fait qu'il est pratiquement impossible de contrôler le matériau même qui a permis l'apprentissage de *GPT-3*, ses performances sont néanmoins impressionnantes et ouvrent pour la première fois la voie à des applications d'écriture semi-automatisée dans le domaine de l'art, du marketing et du journalisme.

La course aux modèles linguistiques géants le plus performants est donc pleinement engagée. Il faut pour gagner engranger des données linguistiques toujours plus vastes dans des modèles ayant toujours plus de paramètres à ajuster et bien sûr avoir accès à des infrastructures informatiques géantes pour pouvoir efficacement apprendre à partir des données récoltées.

## Thèse 4 : En rompant avec les dynamiques culturelles d'évolution des langues, le capitalisme linguistique est susceptible d'être à la base d'une transformation linguistique majeure.

L'arrivée de modèles capables de produire des textes de grande qualité et de multiplier les services linguistiques est susceptible de produire un certain nombre de transformations linguistiques importantes. Pour plus de clarté, il convient de distinguer les ressources linguistiques **primaires** produites sans médiation algorithmique (conversations écrites, contenus de livres numérisés, etc.) des ressources linguistiques **secondaires** produites comme transformations algorithmiques de ressources primaires.

Par exemple, les algorithmes de traduction automatique permettent de transformer le capital linguistique d'une langue en une autre. Ils ont une double importance économique, car ils proposent des services linguistiques précieux et ils décuplent le capital linguistique accumulé. Mais cette intermédiation algorithmique n'est pas sans effet. De plus en plus souvent, de telles ressources secondaires sont indistinguables des ressources primaires.

Le rôle des bots sur Wikipédia est depuis les origines de l'encyclopédie en ligne d'une grande importance[13]. Plus une langue a peu de contenu sur la plateforme, plus la part des algorithmes dans la production des pages est importante[14]. Elle est, par exemple, pour l'Anglais de 9%, mais peut atteindre 67% pour l'Arménien ou 87% pour le Serbo-Croate.

---

[13] Geiger, R. S. (2011). The Lives of Bots. In Critical Point of View: A Wikipedia Reader (pp. 78–93). Institute of Networked Cultures, Amsterdam. http://papers.ssrn.com/sol3/Delivery.cfm?abstractid=2075015#page=41

[14] Kaplan, F., & Nova, N. (2016). Le miracle Wikipedia. PPUR.

Les algorithmes ne se limitent pas à des tâches de gestions neutres. Certains algorithmes sont explicitement utilisés pour corriger des “biais” de représentation. Par exemple, *Lsjbot* est un bot Wikipedia lancé en 2012 par le contributeur suédois Sverker Johansson auteur de 2.7M articles sur Wikipedia, soit 10 000 par jour. Un des objectifs poursuivis par Johansson avec ce bot est de corriger le déséquilibre des informations présentes sur internet[15].

Les algorithmes qui analysent les langues font difficilement la différence entre ces ressources secondaires et les ressources primaires. *GPT-3* est essentiellement entraîné sur une copie du web et sur Wikipedia et ces deux environnements sont de plus en plus “pollués” par les ressources secondaires. OpenAI n’ayant pas son propre capital linguistique doit composer avec la présence grandissante de ces ressources secondaires. GPT-3 contribue par les textes qu’il produit à accélérer le mélange entre les formes “naturelles” et les formes médiées de la langue, conduisant à ce que nous pourrions appeler une nouvelle forme de **créolisation** basée sur une rétroaction des formes algorithmiques sur la langue elle-même.

Aujourd’hui, dans la sphère occidentale, seul Google a encore accès à des ressources linguistiques primaires (via ses services de prothèses linguistiques et via le patrimoine linguistique qu’il a numérisé notamment dans le cadre de *Google Books*). Cette position dominante ne l’empêche pas de connaître ses premières crises dans ce domaine. Le départ de Google de Timnit Gebru, spécialiste des ramifications sociales et éthiques de l'intelligence artificielle et la publication d’un article identifiant en particulier le coût écologique des modèles linguistiques géants a suscité une polémique dans la presse s’est fait largement écho. Selon l’article publié quelque mois après le départ de Gebru dans une conférence sur l'équité et la technologie[16], entraîner un modèle linguistique comme le *BERT* de Google serait l’équivalent en termes de production de $CO_2$ de faire un aller-retour entre New York et San Francisco. Gagner marginalement sur une tache particulière des compétitions linguistiques se fait au prix d’une dépense énergétique considérable. Le capitalisme linguistique est un capitalisme énergétique dans la mesure où il fait maintenant massivement appel à l’intelligence artificielle[17].

L’article soulignait aussi le risque représenté par la reproduction des “biais” présents dans les textes d’entraînement (c’est-à-dire possiblement l’ensemble du Web pour *GPT-3*). Malgré la taille et la diversité des données récoltées, la production de textes produits par des “minorités” est potentiellement réduite dans les corpus utilisés. Ainsi, les compétences linguistiques de *GPT-3* seraient influencées par la position dominante des pays riches qui sont les principaux producteurs des textes présents sur le Web. Comme l’annonce son titre *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Sur les dangers des perroquets stochastiques : Les modèles de langage peuvent-ils être trop grands ?)*,

---

[15] « Il y a davantage de contenu sur Wikipédia sur les hommes que sur les femmes ; un village en Europe aura plus d'informations dans son article qu'une grande ville en Chine. » https://tobyskinnersite.wordpress.com/2017/06/11/the-worlds-most-prolific-writer/ (Consulté le 15 Juin 2020)

[16] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜 Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

[17] Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.

l'article dénonce les effets de modèles qui se contentent de reproduire une structure statistique induite, donne l'illusion de produire un discours et empêche potentiellement le progrès social. Dans la lignée de cette polémique, l'étude des "biais" et des productions toxiques dans les grands modèles linguistiques a donné lieu à d'autres études[18][19][20].

Google aurait affirmé que ses grands modèles linguistiques étaient "conçues pour éviter" les pièges décrits dans l'article[21] et Open AI travaille à corriger les biais des modèles en créant ce qu'ils appellent un ensemble de données "ciblé sur les valeurs" (Process for Adapting Language Models to Society - PALMS)[22]. L'objectif déclaré est de "détoxifier" les modèles pour les rendre compatibles avec des valeurs sociétales en cours d'explicitation et de négociation.

De manière paradoxale, alors que la question du renforcement des stéréotypes statistiques dans les écrits algorithmiques retient l'attention internationale et est à la base de nombreuses nouvelles recherches et publications, peu fond cas de la transition majeure que représente la production massive de textes qui ne sont plus basés sur un processus d'interactions culturelles distribuées mais principalement sur l'optimisation de critères extrinsèques, comme la compatibilité avec certaines valeurs sociales. Les ressources linguistiques secondaires, désormais massivement produites, ne sont même plus le reflet statistique imparfait des ressources linguistiques primaires, c'est un matériau d'une toute autre nature, globalement modelé soit selon des objectifs commerciaux ou des principes éthiques. La production massive de ce matériel linguistique, immédiatement indistinguable des sources linguistiques primaires provoque une rupture sans précédent de la trajectoire culturelle des langues vivantes. Alors que le phénomène de créolisation suit sont cours, les générations futures liront et écriront une langue dont les dynamiques de production sont totalement inédites.

---

[18] Abid, A., Farooqi, M., & Zou, J. (2021). Persistent Anti-Muslim Bias in Large Language Models. *ArXiv:2101.05783 [Cs]*. http://arxiv.org/abs/2101.05783

[19] Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong, and Stephen Denuyl. 2020. Social Biases in NLP Models as Barriers for Persons with Disabilities. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Online, 5491–5501. https://doi.org/10.18653/v1/2020.acl-main.487

[20] Samuel Gehman, Suchin Gururangan, Maarten Sap, Yejin Choi, and Noah A. Smith. 2020. RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models. In Findings of the Association for Computational Linguistics: EMNLP 2020. Association for Computational Linguistics, Online, 3356–3369. https://doi.org/10.18653/v1/2020.findings-emnlp.301

[21] Simonite, T. (2021). What Really Happened When Google Ousted Timnit Gebru. Wired. Retrieved June 17, 2021, from https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened

[22] Solaiman, I., & Dennison, C. (2021). Process for Adapting Language Models to Society (PALMS) with Values-Targeted Datasets. Retrieved online : https://cdn.openai.com/palms.pdf

## Thèse 5 :Les modèles linguistiques conduisent à des services d'expression performatifs optimaux adaptés aux contextes et intentions des actes de langage.

La recherche Google, même dans sa version la plus simple était déjà en 1998, un service linguistique qui transformait un énoncé (la requête) en un résultat (une liste de page). C'était un langage de communication vers la machine plus souple que l'antique ligne de commande[23] mais moins puissant que la langue naturelle : un idiome adapté aux contraintes d'interprétation des interfaces homme-machines, mais néanmoins motivé par la pertinence et l'effectivité.

En vingt ans, les progrès des modèles linguistiques ont permis des avancées impressionnantes en termes de compréhension et surtout en performativité. La parole est devenue interface d'action pour choisir parmi des catalogues gigantesques d'objets, programmer des appareils simples de domotique, informer ses proches, rechercher un itinéraire. Avec la maturité des technologies de traduction basée en grande partie sur la révolution des *Transformers*, la communication professionnelle dans des langues inconnues est devenue possible en pratique dans des contextes simples, mais aussi l'écriture semi-automatisée de programmes informatiques simplement sur la base du résultat attendu.

Les actes de langage, jadis un modèle d'interprétation du phénomène linguistique sont devenus des routines d'action concrètes dans le monde numérique. Par la médiation algorithmique, la parole permet d'agir sur monde. L'acte de langage commence par une intention ou un but et se déploie vers un effet qui aura ou n'aura pas lieu [24]. La réussite ou l'échec de l'acte dépend en grande partie des compétences du locuteur de son habilité à structurer ses phrases pour obtenir l'effet attendu.

La médiation algorithmique permet l'optimisation de cette transformation de l'intention vers l'effet attendu par plusieurs moyens. D'abord, en donnant au locuteur la possibilité de puiser dans des modèles entraînés sur des volumes linguistiques immenses facilement déployables par les interfaces de l'écriture régulée. Dès qu'il entame le passage à l'acte linguistique, des arborescences de possibilités s'ouvrent à lui. Ensuite en lui permettant de contrôler son énonciation par l'intermédiaire de paramètres de plus haut niveau, choisissant selon son objectif la forme plus optimale. Le locuteur pourra, par exemple, choisir entre une écriture de style formel ou informelle, ou, pourquoi pas, l'assurance d'une langue "détoxifiée", garantie sans haine implicite contre les minorités ou biais liés au genre. Comme un outil de composition automatique qui vous permet de produire des mélodies nouvelles en choisissant un ton, une gamme, un style rythmique sans nécessairement que vous ayez à maîtriser le solfège, l'harmonie et la virtuosité de l'instrument, les mots pourront être choisis selon le contexte pour obtenir l'effet optimal avec les nuances adaptées. Leurs effets sur leur interlocuteur, humain ou machine sera alors maximal. Avec l'orchestration algorithmique, les mots s'ajustent au monde, le monde s'ajuste aux mots.

---

[23] Stephenson, N. (1999). In the Beginning...was the Command Line. William Morrow Paperbacks.
[24] John Langshaw Austin, How to Do Things with Words. Cambridge (Mass.) 1962 - Paperback : Harvard University Press, 2e ed., 2005, - en français : Quand dire, c'est faire, Seuil,

Face à ces possibilités nouvelles, il devient clair que la bourse des mots n'était que le premier régime du capitalisme linguistique. Le développement de services d'énonciation sous la forme d'acte de langage optimal ouvre un champ des possibles économiques essentiellement encore vierge. Ce second régime du capitalisme linguistique, anticipé par Pascal Chabot dans une pièce de théâtre[25], consiste à vendre l'usage même de la langue. Le jour où les machines nous aideront à écrire mieux est déjà arrivé et avec lui la possibilité de proposer une transformation de la propre locution simplement en payant le service correspondant. Paierons-nous à la phrase ? Sera-t-il possible de souscrire un abonnement ? Le service sera-t-il d'abord gratuit, puis payant pour les locuteurs les plus demandeurs ? En cas de difficulté financière, faudra-t-il compter ses mots ?

## Thèse 6 : La généralisation des prothèses pour l'écriture régulée pourrait conduire, à terme, à une atrophie expressive et prolétarisation des compétences linguistiques

Les prothèses sont le prolongement de notre cerveau. Elles permettent d'agir avec plus de force et d'efficacité dans le monde, mais, irrémédiablement, elles conduisent simultanément à l'atrophie d'une compétence préalable, devenue obsolète ou tout du moins sous-optimale. Il est donc raisonnable de se poser la question d'une possible atrophie expressive résultant de la génération des médiations algorithmique pour toutes les fonctions qui relèvent de l'action par le langage. Alors que le geste langagier médié devient machinique que resta-t-il de nos compétences linguistiques primaires, de moins en moins exercées par manque d'usage ? De la même manière que la généralisation de l'expression par clavier interposé a érodé notre capacité pour l'écriture manuscrite et provoque indirectement des transformations cognitives potentiellement significatives, le développement de la méditation algorithmique entre l'intention et l'expression et de l'écriture régulée peuvent provoquer un déplacement des compétences linguistiques. Nous saurons optimalement diriger notre expression vers un but, mais aurons oublié comment composer un discours seul sans intermédiation technique.

Cet effet secondaire du développement du capitalisme linguistique n'est que peu surprenant. Comment le capitalisme linguistique ne conduirait-il pas une **prolétarisation**, c'est-à-dire une perte des savoirs[26] ? La prothèse linguistique automatise le savoir linguistique collectif et aliène potentiellement par ce biais le locuteur, même si elle lui donne une effectivité performative plus grande dans le monde. La langue qui se transmettait et se négociait de génération en génération, **la langue artisanale,** est aujourd'hui intégrable et manipulable dans un colossal hyperobjet numérique[27]. Cette rupture conduit à des nouvelles possibilités

[25] Chabot, P. (2018). L'homme qui voulait acheter le langage. Presses Universitaires de France. https://doi.org/10.3917/puf.chabo.2018.01
[26] Voir entrée "Proleterisation" dans Stiegler, B. (2013). *Pharmacologie du Front national: Suivi du vocabulaire d'ars indistrualis*. Flammarion.
[27] Morton, T. (2013). Hyperobjects: Philosophy and Ecology After the End of the World. University of Minnesota Press.

expressives et à la progressive disparition des compétences anciennes, celles d'avant la médiation algorithmique.

Cet effet d'atrophie sera-t-il confiné à l'expression écrite ? Les prothèses linguistiques remontant la chaîne des actes de langage pour se rapprocher toujours plus de l'intention et prenant en charge la production linguistique, la différence entre écrit et oral s'atténue. La même intention communicative peut donner lieu à une production médiée d'écrit ou d'oralité, notamment dans le cadre des échanges sur les plateformes numériques. Ce qui freine leur déploiement sur l'oralité est aujourd'hui uniquement la difficulté à proposer des interfaces d'écriture régulée par oral. Dans les domaines de l'expression où le contrôle arborescent n'est pas utile, comme les ordres de programmations simples, elles sont déjà à l'œuvre.

L'oralité médiée par la machine est évidemment différente de l'oralité artisanale. C'est une oralité appareillée, amplifiée dans la mesure où elle permet l'action et la commande, mais séparée de son milieu originel, social et réactif, le plus souvent improvisé : une parole désormais faite de formules magiques plus de que conversations.

En une vingtaine d'années, le capitalisme linguistique en créant d'abord de la valeur en interceptant et revendant l'accès aux flux linguistiques a pu produire un changement d'échelle réifiant pour la première fois l'usage de la langue sous la forme de modèles géants. Alors que nous découvrons progressivement la puissance expressive de ces modèles et que nous comprenons ce que contient vraiment le savoir linguistique une fois capturé dans les filets du numérique, alors que les premières expériences d'optimisation globales de la langue sont poursuivis à la recherche d'une langue neutralisée, sans histoire, quasi-adamique[28], nous assistons presque passifs à une transformation culturelle majeure, sans doute un point de non-retour linguistique, qui déterminera le destin des mots et des idées pour les décennies à venir.

Alors qu'avec l'aide des machines, nous parlerons et écrirons mieux, nous perdrons peut-être d'abord la capacité puis le goût de nous exprimer. Si le capitalisme linguistique réussit à transformer la nature profonde de son d'objet d'exploitation, si de créoles humain-machines les langues finissent simplement comme des systèmes globaux à optimiser pour satisfaire des fonctions de coût sociétales, si un jour « résister c'est se taire » mais alors peut-être, conscients que, depuis la tour de Babel, les langues n'ont eu cesse de diviser les hommes, nous trouverons la paix et l'harmonie dans un nouveau mutisme. La parole artisanale n'aura alors été qu'une parenthèse technologique.

---

[28] Eco, U. (1994). La recherche de la langue parfaite dans la culture européenne. Éditions du Seuil.