

Stylistique de la recherche linguistique en IA: de LISP à GPT-3

Pierre Depaz

Paris-3 Sorbonne Nouvelle - THALIM

1. Introduction

En Septembre 2020, le quotidien britannique The Guardian publie un éditorial entièrement généré par une intelligence artificielle (GPT-3). Cette troisième génération du modèle *Generative Pre-trained Transformer* (GPT-3), rendue publique en Mai 2020, a représenté un événement significatif, de par sa capacité inédite à générer des textes de qualité similaire à ceux produits par des humains, et souvent indiscernables de ces derniers.

Malgré ces succès récents, la recherche dans le domaine du *traitement automatique des langues* (TAL), champ de l'informatique se concentrant sur la compréhension et l'expression en langues humaines (par contraste aux langages de programmation), est pourtant concomitante des débuts de la recherche en Intelligence Artificielle (IA) au début des années 1950. Au vu des réalisations de GPT-2, GPT-3 et des autres modèles basés sur l'architecture Transformer, il semble cependant que nous soyons, à première vue, confrontés à un saut qualitatif, plutôt que purement quantitatif en termes de génération de texte. Quels sont les ressorts théoriques et pratiques de ce saut? Et comment ces ressorts peuvent-ils nous présenter une nouvelle perspective sur les textes qui nous entourent, de la littérature aux messages de forums en passant par des manuels techniques?

Plutôt que présenter une histoire strictement rationnelle des développements techniques et des objets mathématiques sur lesquelles sont basés les modèles Transformer, il s'agira ici d'identifier une *stylistique* de la recherche en intelligence artificielle, et d'élucider, par comparaison de deux stylistiques différentes, un mouvement dialectique dans l'appréhension du langage en général, et de l'écrit en spécifique.

Après avoir élucidé la notion de style que nous utiliserons, en tant que liaison entre manière de penser et manière de faire, nous examinerons les influences théoriques, pratiques notationnelles et implémentations concrètes des deux courants principaux du TAL: l'approche symbolique, des années 1950 aux années 1970, et l'approche statistique, des années 1950 aux années 2020. Finalement, nous montrerons que l'approche

statistique des langues implique une conception matérielle et contiguë de des œuvres textuelles qui, de ce fait, contribue à une redéfinition par extension du domaine de la littérature.

2. La notion de style en sciences

Afin de capturer les aspects anthropologiques et sociaux implicites à la recherche et création scientifique (Collins and Pinch, 1998; Sack, 2019), cette communication considère donc les processus qui ont abouti à la réalisation des modèles Transformer d'IA à travers la notion de style, développée notamment par Gilles Gaston-Granger (Granger, 1988) et Ian Hacking (Hacking, 1994). Il ne s'agit pas ici de réduire cette notion à un ensemble cohérent de manifestations esthétiques et de représentations formelles (Preziosi, 2009) ou à un système linguistique (Sebeok, 1966), mais de connecter ces manifestations et ces signes à une posture épistémologique, à une manière d'appréhender un ensemble abstrait de problèmes théoriques.

En effet, le travail de Gaston-Granger cherche à mettre en exergue comment les différentes manières de faire des mathématiques, d'Euclide à Descartes, réorganise la primauté de certains concepts (e.g. la grandeur, la distance, l'angle, le vecteur) par rapport à d'autres, au sein de la recherche en géométrie. Cette approche trouve un écho dans le travail des études en sciences et technologies, lorsqu'il s'agit de mettre en évidence l'impact des réalités concrètes, techniques et sociales, de la recherche scientifique et le rôle joué par les dispositifs spécifiques déployés lors de cette recherche dans les résultats et innovations techniques qui en découlent (Star, 1992; Latour, 1987). Nous partons donc d'une interaction rapprochée entre contextes historiques et théoriques, buts avérés de la recherche scientifique, et des outils déployés afin d'atteindre ces buts—une interaction qui nous permettra de mieux saisir les enjeux d'un langage généré par une IA.

Ces contextes historiques et théoriques seront appréhendés en tant que réseau de références explicites et implicites, parfois identifiés en tant que *tradition* au sein de laquelle s'insère le travail de chercheurs. Ces derniers sont constitués de références académiques (paradigmes présents et passés), fictionnelles (œuvres narratives et poétiques évoquées par les acteurs de cette recherche), économiques (influence de la configuration du marché à un moment donné) et techniques (possibilités concrètes d'implémenter une idée donnée). Plus spécifiquement, nous considérons ici les œuvres fictionnelles comme le fruit d'un contexte, d'une intention, d'une tradition culturelle, d'une réception du lectorat et de conventions

stylistiques(Genette). Si des textes religieux, des traités de politiques et des épiques nationaux existent aussi bien en tant qu'ouvrages de fictions et ouvrages pratiques, nous pouvons également considérer les ouvrages scientifiques comme tels.

Les buts avérés de la recherche scientifique sont, eux, manifestés principalement par des ensembles d'évaluations (ensembles de problèmes à résoudre, tests, *benchmark*, formulaires de questions, etc.) établissant de manière implicite l'envergure de la recherche elle-même. De la même manière que les tests de quotient intellectuel (QI) ne prennent en compte qu'une certaine facette de l'intelligence humaine et ne sont utiles que dans certains contextes (Weinberg, 1989), les cadres d'évaluations des technologies d'IA vont influencer ce que les implémentations concrètes de ces IA vont pouvoir accomplir.

Enfin, les outils déployés seront ici considérés en deux catégories—les outils notionnels et les outils computationnels—en tant que manifestations concrètes de la pensée, mais aussi en tant qu'acteurs permettant ou obstruant un déploiement de cette pensée dans le monde réel. Les outils notationnels, comme nous le verrons, sont ceux qui permettent une *transcription* et une *communication* de la pensée (tels que la liste, le symbole, le graphe). Les outils computationnels sont ceux qui permettent une *activation* de la pensée, et consistent donc en programmes et en langages de programmations, produisant des résultats concrets. Dans le contexte de cette communication, nous nous intéresserons principalement aux langages de programmation LISP et Python, implémentations informatiques des approches symboliques et statistiques du langage, respectivement.

3 Deux styles de recherche en IA

L'approche symbolique du langage en TAL est concomitante de la fondation du champ de recherche en IA elle-même. La conférence de recherche de Dartmouth, se déroulant en 1956 et organisée par John McCarthy, Marvin Minsky et Nathaniel Rochester, avec le support de Claude Shannon, se fixe alors pour but de déterminer la manière dont “les machines utilisent la langue, forment des abstractions et des concepts, résolvent le genre de problème à présent réservés aux humains, et s'améliorent”(McCarthy, Minsky, et al.). La posture épistémologique dominante est alors celle d'une approche symbolique, tant par son contexte scientifique et historique que par les outils programmatiques déployés.

3.1 L'approche symbolique du langage

Cette tradition scientifique trouve tout d'abord son origine dans les œuvres de Gottfried Leibniz et de Gottlob Frege, et particulièrement en ce qu'elles furent redécouvertes par Bertrand Russell et Ludwig Wittgenstein au début du XXe siècle (Williams, 2002). Alan Turing fréquentera ce dernier à Cambridge, après avoir publié l'article fondateur du champ de l'informatique (Floyd, 2019). Les concepts établis au sein de cette généalogie de pensée se cristallisent autour d'un formalisme de logique mathématique, et de génération de sens à travers une manipulation de symboles. Ces paradigmes de pensée sont illustrés principalement par des styles de notation favorisant largement *le symbole* et *la liste*, qui apparaissent ensuite, comme éléments centraux des premiers outils d'IA, notamment au sein du langage LISP.

Le symbole, en tant qu'élément conceptuel et clairement délimité, occupe une place prépondérante dans l'œuvre *De Arte Combinatoria*, de Leibniz. Partant du postulat que "toute chose est, ou n'est pas", il tente d'utiliser cette approche qui sera *a posteriori* celle de Georges Boole, afin de manipuler chacune de ces choses (concepts) de manière formelle, à travers une *characteristica universalis*. Les possibilités d'un tel système de manipulation formelle de symboles permettrait donc de déterminer n'importe quel concept ou abstraction via une manipulation logico-mathématique:

If we had it [a characteristica universalis], we should be able to reason in metaphysics and morals in much the same way as in geometry and analysis... If controversies were to arise, there would be no more need of disputation between two philosophers than between two accountants [...] Let us calculate. (Russell, 1950)

Le symbole en tant que dispositif stylistique permettrait donc la (re)combinaison de simples éléments primaires (appelés monades chez Leibniz) en des expressions complexes. Ce dispositif stylistique de notation se retrouvera notamment dans les notations utilisées par Georges Boole et Gottlob Frege; à chaque fois, il s'agit de poser qu'il est possible de capturer, par le biais d'une représentation formelle, *l'essence* d'un concept, et de le rendre manipulable à travers sa stricte délimitation—postulat-clé du travail de Turing lorsqu'il décrit sa machine.

À cet effet d'unicité du symbole vient s'ajouter un second mécanisme stylistique—celui de la liste. Lorsque le travail de Leibniz est réinterprété par Russell, ce postulat d'unicité est requalifié par les calculs des prédicats

modernes, notamment au travers la notation Peano-Russell. Il s'agit alors de considérer autant la relation entre éléments primaires que ces éléments eux-mêmes, en introduisant donc la nécessité d'établir un *lien* entre ces symboles, par le biais de la liste.

La liste est le style de notation dominant dans des ouvrages tels que le *De Arte Combinatoria* de Leibniz, le *Begriffsschrift* de Frege, le *Principia Mathematica* de Russell ou le *Tractatus* de Wittgenstein, et, tout comme le symbole, implique une certaine posture épistémologique. La liste extrait ses éléments de leur existence située d'origine, et les reconnecte de manière rigoureusement définie. La constitution d'une liste est donc l'exemple de la sorte de décontextualisation que l'écrit promeut et qui, selon Goody, permet à l'esprit d'exercer une certaine emprise sur la "réalité" (Goody, 1977).

Présente dès les premiers inventaires, manuels, édits administratifs, la liste est une manière de prendre des symboles, des éléments pictographiques, sémantiquement cohérents, afin de les ré-assembler et de re-constituer le monde. La liste établit des limites claires; elle est simple, elle est discontinue et abstraite.

La liste et le symbole ne sont pour l'instant que des systèmes stylistiques d'appréhension et de construction du monde, un monde considéré comme séparable, puis reconstituable. Cette approche encore principalement conceptuelle se voit alors activée à travers des outils de programmation. Particulièrement, les langages IPL et LISP, ainsi que le programme SHRDLU vont permettre l'implémentation concrète de ces systèmes de notations afin d'élucider les problèmes posés lors de la conférence de Dartmouth, et de préciser par la-même les buts avérés de cette recherche.

3.2 Les langages symboliques

Au même moment que la conférence de Dartmouth sont développés deux langages qui mettent concrétisent une approche symbolique du langage. D'une part, IPL (*Information Processing Language*) conçu par Allen Newell, Cliff Shaw et Herbert A. Simon, en 1956; de l'autre LISP (*LISt Processor*), conçu par John McCarthy, organisateur de la conférence, en 1958.

L'élément de base d'IPL est le *symbole*, en tant qu'unité signifiante qui est directement relié aux différentes locations de mémoire du matériel informatique sur lequel s'exécute le langage. Le symbole représente donc un

espace conceptuel qui peut contenir n'importe quelle valeur, tant qu'elle est conceptuellement distincte et disjointe d'autres valeurs (Newell and Tonge, 1960). Cette architecture correspond particulièrement aux genres de problèmes qu'IPL permet de résoudre. En effet, il est initialement conçu pour démontrer les théorèmes exposés par Russell dans ses *Principia Mathematica*, mais est également utilisé afin d'écrire des programmes d'IA, tels que le *General Problem Solver*, le *Logical Theorist*, ainsi que quelques programmes de jeu d'échecs et de dames (Samuel, 1960). Il s'agit ici de problèmes hautement formalisés, basés sur la pré-existence de certaines règles formelles, avec des données de départ qui se calquent facilement sur le fonctionnement symbolique d'IPL—postulats mathématiques, propositions logiques, et positions de pièces de jeu.

Lorsque LISP remplace IPL comme langage de l'IA quelques années plus tard, il hérite non seulement de cette approche symbolique, mais la compose également avec la liste. Cette dernière est l'élément de base du langage et est considérée elle-même comme un symbole. En introduisant la manipulation récursive de ces listes, LISP met en exergue une manipulation des données atomistique, puis relationnelle—chaque élément peut être appréhendé par soi-même, ou par sa relation à d'autres éléments mitoyens (McCarthy, Levin, et al., 1965). Cette approche du problème, particulièrement adaptée à des structures en arborescence, fait écho au travail de Noam Chomsky, contemporain de MacCarthy au Massachusetts Institute of Technology. Chomsky vient alors de présenter une théorie de la structure du langage en tant qu'arborescence (Chomsky, 1957), décomposant une phrase en unités cohérentes reliées entre elles.

Cette conception symbolique du langage se trouve alors appliquée à la même catégorie de problèmes hautement formalisés qu'IPL, jusqu'au développement de SHRDLU en 1978. Conçu dans un dialecte de LISP, MacLisp par Terry Winograd, SHRDLU confronte une approche symbolique du monde avec sa réalité tangible. SHRDLU a pour but de démontrer la compréhension d'un langage naturel dans le contexte d'un environnement physique hautement symbolisé, présupposant la compréhension du sujet afin d'interagir avec un langage naturel (Winograd, 1971). Pourtant, même dans ce genre d'environnement, l'IA selon LISP ne remplit pas ses promesses.

Pourtant, l'approche symbolique de LISP se retrouve confrontée à deux obstacles dans les années 1980: un matériel informatique limité et le fait qu'elle se soit cantonnée à des *toy problems* (problèmes enfantins, d'après l'expression de Seymour Papert) (Nilsson, 2009), tels que des jeux et théorèmes aux règles bien définies. Réduire le monde à une série de symboles manipulables logiquement s'avère être insuffisant pour saisir les

nuances d'un langage qui, lors de situations réelles, s'avère bien plus complexe qu'un traité de mathématiques; un revers qui fait écho à celui de Wittgenstein dans ses *Recherches Philosophiques*. La réponse à cette limitation du symbolisme viendra d'un autre courant, celui de la statistique matérialiste.

3.3 L'approche statistique du langage

Un autre participant à la conférence de Dartmouth était Warren McCulloch, neurophysiologue et cybernéticien. Venant d'une tradition et narrative et biologique, nous allons voir comment McCulloch et ses collaborateurs vont développer un style notationnel diagrammatique qui fait apparaître les notations de statistique, de spatialité, et de plasticité, dont l'implémentation sera complétée par des langages tels que Python et C++.

L'environnement fictionnel dans lequel il évolue est alors principalement composé d'une part des théories de Sigmund Freud, et de l'autre de celles de Norbert Wiener. Lors d'une série d'interventions au Chicago Literary Club, McCulloch va démontrer les limitations d'une psychologie humaine basée sur une approche loin du style scientifique auquel il adhère, une approche psychanalytique où "*les données, les méthodes et les théories sont indissolubles*", et dont "*la dépendance des données sur la théorie [la] sépare de toutes les autres véritables investigations*" (McCulloch, 1953). Si l'interprétation narrative du fonctionnement du cerveau humain n'est pas satisfaisante, il propose alors une approche matérielle, biologique, combinée à la communication stochastique de Wiener.

Cette stochastique dépend d'une situation donnée afin de prévoir la probabilité d'une situation future, une approche empirique et déductive du problème qui contraste avec l'approche théorique et inductive de la tradition de Leibniz (Wiener, 1961). Le travail de Wiener sur les filtres et le mouvement de Brown se voit représenté par une utilisation prépondérante des *intégrations* mathématiques. Ces dernières sont une notation mathématique particulière qui prend en compte *tous les événements d'une période donnée*, plutôt que de les séparer en symboles distincts (Wiener, 1930, 1933).

Dans ce contexte d'une approche plastique et stochastique, McCulloch tente, dès 1943, de démontrer avec Walter Pitts que le cerveau est une machine de Turing et que son unité de base n'est pas le symbole, mais le neurone (McCulloch and Pitts, 1943). Ils y combinent le style de la logique prédicative de Russell avec

l'approche systématique et empirique (*input/output*) de la cybernétique et représentent les interactions complexes de différentes parties d'un système de manière *diagrammatique*. Ce style de notation, qui remonte à Euler et à sa théorie de graphes, montre qu'il est possible d'appréhender un problème aux fondations mathématiques à travers une représentation graphique qui ne soit pas du ressort de la géométrie. D'après Sybille Krämer, les inscriptions diagrammatiques sont des media qui permettent un point de connexion entre la pensée et l'intuition, entre le noétique et l'esthétique. Des entités théoriques peuvent être spatialisées, être rendues intuitives et manipulables (Krämer, 2013). La différence avec les diagrammes arborescents de Chomsky se situe au niveau où le style utilisé par McCulloch permet de prendre plus facilement en compte les stimuli et sources d'informations externes au système décrit.

Dans ce contexte, le langage est considéré comme un objet stochastique, où la compréhension se passe au niveau même de la matière utilisée (Lettvin et al., 1959), où les séquences de mots utilisées en langage naturel nous semblent si familières que, en ayant entendu une partie, nous établissons la suite correctement, n'y prêtant qu'une attention distraite (McCulloch, 1953).

3.4 Des langages protéiformes pour une langue spatialisée

Ce style diagrammatique se voit directement traduit en des projets d'IA dès le début des années 1950. Le SNARC de Marvin Minsky et Dean Edmonds en 1951 (Crevier, 1993) et le Mark I Perceptron de Frank Rosenblatt en 1960. Ces deux projets sont plus matière que logiciel, consistant en des enchevêtrements de câbles, agissant eux-même comme neurones artificiels. Il s'agit alors de prendre en compte la totalité d'une situation, contrairement à son essentialisation par des moyens symboliques, et d'estimer la conclusion la plus probable, plutôt que celle qui serait mathématiquement vérifiable, cette fois testés sur des standards tels que la sortie d'un labyrinthe.

Ici encore, ces premières tentatives se heurtent aux limitations du matériel informatique de l'époque, mais c'est cette approche du langage liée à la neuroplasticité qui prédominera lors de l'émergence de l'apprentissage profond au début des années 2010 qui, après un "hiver de l'IA", font la part belle à des modèles probabilistes, à la combinaison de divers algorithmes et à la supervision humaine des résultats (Nilsson, 2009). Cette approche spatiale de l'intelligence et du langage se retrouve alors traduite dans les langages utilisés afin

d'implémenter ces réseaux neuronaux.

Le langage Python devient le langage de prédilection de cette nouvelle IA, dont sont issus les modèles *Transformer*, dont GPT-3 fait partie. Par rapport à LISP, on peut identifier une perte de rigueur mathématique pour un gain d'effectivité, laissant les concepts linguistiques se définir *de facto* et *in situ*.

Tandis que LISP se limite à une seule unité de base, Python permet une interface facile de plusieurs types de données de manière quasiment simultanée (une particularité des *scripting languages* dont Python fait partie). Plus particulièrement, Python permet l'utilisation de *data-frames*, ensemble de données manipulées simultanément. Ces *data-frames* conviennent particulièrement à une représentation du langage en tant que vecteurs dans un espace à un très grand nombre de dimensions (Mikolov et al., 2013). Cette représentation vectorielle implique une continuité du langage qui demeurerait inexistante, ou du moins compliquée à représenter, dans une conception symbolique des langues naturelles. Un mot, une proposition, une phrase ou un texte sont représentés par un point et une direction, et donc par une relation spatiale à tous les autres mots, propositions, phrases ou textes pouvant potentiellement exister. Cette relation est immédiate et empirique, plutôt qu'une parenté abstraite qui demande la traversée entière d'une structure arborescente afin de connecter deux éléments linguistiques.

Le style des cybernéticiens faisant la part belle à la communication entre organismes incarnés, Python permet cela en permettant l'utilisation optimale des différents matériaux qui composent un ordinateur. Particulièrement, son implémentation en C permet une communication extrêmement rapide avec le CPU et le GPU, deux unités de calculs—respectivement séquentielle et parallèle. Nous retrouvons aussi, d'une certaine manière, la nécessité d'une incarnation et d'une situation favorisée par les linguistiques statistiques.

C'est à partir de l'environnement fictionnel des recherches en statistique et neurophysiologie, telle que l'idée que le cerveau est avant tout un organe matériel et biologique, et qu'il est également une machine de Turing, par d'un style de notation qui permet l'interconnexion and l'incorporation de données externes, et au travers d'outils pratiques qui permettent la manipulation rapide (et donc matérielle) d'une représentation spatiale du langage, que se trouve la conception et le développement du modèle GPT-3. Ce postulat du langage comme spatial et matériel comporte des implications pour notre conception de l'écrit et du littéraire vers lesquelles nous nous tournons ensuite.

4 Le langage matériel

4.1 GPT-3

En conséquence d'un changement dans la recherche—changement stylistique, plutôt que fondamental—, et avec l'implémentation Python de GPT-3 par OpenAI, il ne s'agit plus de considérer le langage comme une liste atomique d'entités connectées par des liens syntactiques préétablis, mais au contraire comme un espace vectoriel continu à hautes dimensions. Les textes sont représentés comme donnée spatiale, et accédés comme donnée spatiale, avant d'être convertis à nouveau en format lisible de phrase naturel. Au lieu d'assumer une structure préexistante, telle une grammaire générative, il s'agit de construire un paysage textuel à partir de toutes les relations possibles entre chaque termes extrait du texte.

L'architecture Transformer, sur laquelle GPT-3 est donc basée, crée donc un réseau multidirectionnel de relations entre termes extraits d'un corpus gigantesque. Le mécanisme-clé du système est représenté par un *diagramme*, plutôt que par une liste de symboles: la compréhension et génération du langage passe bien par un style statistique et systématique, plutôt que symbolique (Vaswani et al., 2017). Il est désormais possible d'explorer ce terrain linguistique, comme le font déjà certains projets expérimentaux, notamment *Compasses* d'Allison Parrish, *1 the road* de Ross Goodwin, ou le jeu *AI Dungeon* par Latitude.

4.2 Une littérature étendue

Quelles sont les conséquences d'un texte généré automatiquement, et dont la qualité égale celle d'un auteur humain? La première est celle d'une reconception des barrières de genre. Tandis qu'une approche symbolique se focalisent sur des catégories de texte spécifiques (problèmes de logique, analyse de littérature de jeunesse, conversation dans un environnement contrôlé), l'approche statistique de GPT-3 n'établit pas de distinction entre types de texte; au contraire, c'est la quantité de matière textuelle qui prime¹. Si la définition du genre n'est pas théorique, elle est néanmoins pratique: statistiquement, certaines phrases vont ensemble, d'autres n'apparaissent jamais dans un même texte.

¹ GPT-3 est entraîné sur 175 milliards de paramètres, dix fois plus importants que le deuxième plus grand modèle, Turing NLG.

En ce qui concerne la question de la créativité d'une IA de ce type, un style de la statistique nécessite forcément un corpus préexistant, ce qui implique alors un résultat reposant exclusivement sur la recombinaison de données passées. De par cette approche, la question de la construction artificielle d'un raisonnement, et donc possiblement de création nouvelle, trouve une réponse limitée (Marcus and Davis, 2020; Bender et al., 2021). Puisqu'ils sont extraits d'un ensemble donné, les textes de GPT-3 ressemblent davantage à un poème Dada porteur de sens et de cohérence qu'à un texte communiquant une idée.

Enfin, l'approche statistique nous montre la différence qui existe entre lire et écrire. Il est effectivement possible de générer des articles de presse, ou des romans entiers, sans pour autant pouvoir commenter le contenu de ces articles par le même algorithme. Les buts avérés de cette recherche indiquent que ces deux domaines sont bien disjoints et ne sont pas le point fort de GPT-3.

5 Conclusion

L'émergence de GPT-3 peut donc être considérée comme le fruit d'un changement de style dans la recherche scientifique. Autant l'approche symbolique que l'approche statistique existent au sein d'environnements fictionnels, de buts avérés, de notations et d'outils différents, qui manifestent différentes conceptions du langage. Dépassant la conception strictement symbolique du langage qui trouve un écho dans le domaine de la linguistique générative, l'approche statistique considère davantage le langage comme une matière existant dans l'espace, dont la représentation conditionne alors la manipulation, afin d'obtenir des résultats—certes excellents—dans un domaine bien particulier. Si les avancées technologiques nous proposent de questionner nos conceptions actuelles du langage et du texte, notamment sur les questions de délimitation des genres, et du rôle des textes passés comme influence des textes futurs, c'est aussi l'influence d'un style particulier sur les réalisations techniques contemporaines qu'il faut retenir afin d'en mesurer les opportunités et les limitations.

Bibliographie

Bender, Emily M., et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2021, pp. 610–23. *ACM Digital Library*, <https://doi.org/10.1145/3442188.3445922>.

Chomsky, Noam. "Syntactic Structures." *Syntactic Structures*, De Gruyter Mouton, 1957. <https://doi.org/10.1515/9783112316009>.

Collins, H. M. (Harry M.), and T. J. (Trevor J.). Pinch. *The Golem : What You Should Know about Science*. Cambridge [England]; New York, NY: Cambridge University Press, 1998. *Internet Archive*, <http://archive.org/details/golemwhatyoushou00coll>.

Crevier, Daniel. *Ai: The Tumultuous History Of The Search For Artificial Intelligence*. Basic Books, 1993.

Floyd, Juliet. "Wittgenstein and Turing." *Philosophy of Logic and Mathematics: Proceedings of the 41st International Ludwig Wittgenstein Symposium*, De Gruyter, 2019, pp. 263–96. www.degruyter.com, <https://doi.org/10.1515/9783110657883-016>.

Genette, Gérard. *Fiction & Diction*. Cornell University Press, 1993.

Goody, Jack. *The Domestication of the Savage Mind*. Cambridge University Press, 1977.

GPT-3. "A Robot Wrote This Entire Article. Are You Scared yet, Human? | GPT-3." *The Guardian*, 8 Sept. 2020, <http://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>.

Granger, Gilles-Gaston. *Essai d'une philosophie du style*. 0 edition, JACOB, 1988.

Hacking, Ian. "Styles of Scientific Thinking or Reasoning: A New Analytical Tool for Historians and Philosophers of the Sciences." *Trends in the Historiography of Science*, edited by Kostas Gavroglu et al., Springer Netherlands, 1994, pp. 31–48. *Springer Link*, https://doi.org/10.1007/978-94-017-3596-4_3.

Krämer, Sybille. "'The Mind's Eye': Visualizing the Non-Visual and the 'Epistemology of the Line.'" *Volume 2* *Volume 2*, De Gruyter, 2013, pp. 275–94. www.degruyter.com, <https://doi.org/10.1515/9783110330496.275>.

Latour, Bruno. *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press, 1987.

Lettvin, J. Y., et al. "What the Frog's Eye Tells the Frog's Brain." *Proceedings of the IRE*, vol. 47, no. 11, Nov.

1959, pp. 1940–51. *IEEE Xplore*, <https://doi.org/10.1109/JRPROC.1959.287207>.

Marcus, Gary, and Ernest Davis. “GPT-3, Bloviator: OpenAI’s Language Generator Has No Idea What It’s Talking about.” *MIT Technology Review*, 20 Aug. 2020, <https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/>.

McCarthy, John, Marvin Minsky, et al. *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Self-published, Aug. 1955, <https://web.archive.org/web/20070826230310/http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

McCarthy, John, Michael I. Levin, et al. *LISP 1.5 Programmer’s Manual*. MIT Press, 1965.

McCulloch, W. “The Past of a Delusion.” *Chicago Literary Club*, 1953, <https://web.archive.org/web/20080509163314/http://www.chilit.org/Published%20Papers/56.%20McCulloch%20-%20Delusion.PDF>.

McCulloch, Warren S., and Walter Pitts. “A Logical Calculus of the Ideas Immanent in Nervous Activity.” *Bulletin of Mathematical Biology*, Vol. 5, 1943. pp. 115–133.

Mikolov, Tomas, et al. “Efficient Estimation of Word Representations in Vector Space.” *ArXiv:1301.3781 [Cs]*, Sept. 2013. *arXiv.org*, <http://arxiv.org/abs/1301.3781>.

Newell, A., and F. M. Tonge. “An Introduction to Information Processing Language V.” *Communications of the ACM*, vol. 3, no. 4, Apr. 1960, pp. 205–11. *April 1960*, <https://doi.org/10.1145/367177.367205>.

Nilsson, Nils J. *The Quest for Artificial Intelligence*. Cambridge University Press, 2009.

Preziosi, Donald. *The Art of Art History: A Critical Anthology*. New Edition, Oxford University Press, 2009.

Russell, Bertrand. “Logical Positivism.” *Revue Internationale de Philosophie*, vol. 4, no. 11, *Revue Internationale de Philosophie*, 1950, pp. 3–19.

Sack, Warren. *The Software Arts*. MIT Press, 2019.

Samuel, Arthur L. “Programming Computers to Play Games.” *Advances in Computers*, edited by Franz L. Alt, vol. 1, Elsevier, 1960, pp. 165–92. *ScienceDirect*, [https://doi.org/10.1016/S0065-2458\(08\)60608-7](https://doi.org/10.1016/S0065-2458(08)60608-7).

Sebeok, Thomas A. *Style in Language*. Cambridge (Mass.): Massachusetts Institute of Technology, 1966. *Internet Archive*, <http://archive.org/details/styleinlanguage00sebe>.

Star, Susan Leigh. “Craft vs. Commodity, Mess vs. Transcendence:” *The Right Tools for the Job*, edited by Adele E. Clarke and Joan H. Fujimura, Princeton University Press, 1992, pp. 257–86, <http://www.jstor.org/stable/j.ctt7zv5cv.13>. JSTOR.

Vaswani, Ashish, et al. "Attention Is All You Need." *ArXiv:1706.03762 [Cs]*, 5, Dec. 2017. *arXiv.org*, <http://arxiv.org/abs/1706.03762>.

Weinberg, Richard A. "Intelligence and IQ: Landmark Issues and Great Debates." *American Psychologist*, vol. 44, no. 2, American Psychological Association, 1989, pp. 98–104. *APA PsycNet*, <https://doi.org/10.1037/0003-066X.44.2.98>.

Wiener, Norbert. *Cybernetics Or Control and Communication in the Animal and the Machine*. MIT Press, 1961.

Wiener, Norbert. "Generalized Harmonic Analysis." *Acta Mathematica*, vol. 55, no. none, Institut Mittag-Leffler, Jan. 1930, pp. 117–258. *Project Euclid*, <https://doi.org/10.1007/BF02546511>.

Wiener, Norbert. *The Fourier Integral and Certain of Its Applications*. *Nature* vol. 132, pp. 731–732. 1933.

Williams, Sam. *Arguing A.I. : The Battle for Twenty-First-Century Science*. New York : AtRandom.com, 2002. *Internet Archive*, <http://archive.org/details/arguingai00samw>.

Winograd, Terry. *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*. Jan. 1971. *dspace.mit.edu*, <https://dspace.mit.edu/handle/1721.1/7095>.