

TECHNOLINGUISME ET MULTILINGUISME SUR WEB 3.0-UBEK ET RICA

Yohanna Joseph WALIYA

Le Village Français du Nigéria (Centre Inter-universitaire Nigérian d'Études Françaises), Ajara-Badagry,
Nigéria

E-mail: waliyayohannajoseph@yahoo.fr, ORCID ID: 0000-0003-4456-7458

TECHNOLINGUALISM AND MULTILINGUALISM ON THE WEB 3.0: UBEK ET RICA BLOG

Yohanna Joseph WALIYA

The Nigeria French Village (Inter University Centre for French Studies), Ajara-Badagry, Nigeria

E-mail: waliyayohannajoseph@yahoo.fr, ORCID ID: 0000-0003-4456-7458

ABSTRACT: The democratisation of the Internet has given rise to a digital society, the most influential, smart, and cumbersome in the history of humanity. English, the most used language on the Internet, accounts for 49.7% of the linguistic content across all digital platforms. The dominance of English has posed a linguistic challenge to Francophone, Hispanophone, Russophone, Lusophone, Sinophone, Arabophone, and many other linguistic communities for several decades. English continuously enriches itself with emerging technological buzzwords thanks to linguistic content disseminated on websites and digital social network applications such as ChatGPT, Google, Facebook, Twitter, YouTube, Instagram, etc. This research focuses on the sociolinguistic competition between English and French in the virtual world, analysing code-switching (English and French) on the French weblog Usbek et Rica, which adopts this style to convey information to its Francophone audience. Can it be said that this illustrates the linguistic homogenisation of English as the language of the Internet to the detriment of other languages, such as French? This article aims to answer this question by examining articles published on usbeketrica.com from 2016 to 2020, using these publications as a data corpus. It also critiques the programming languages of the blog in comparison with natural languages in society, applying the Marie-Anne Paveau's ecological linguistic approach, which stipulates that all linguistic and extra-linguistic signs produce symmetrical meanings in the digital ecosystem. The findings show that the impact of English on French, in terms of technological language, is increasingly evident on the blog.

KEYWORDS: Technolinguism, Multilingualism, code-switching, Web 3.0, ecological linguistics, Webology, Usbek et Rica, French, English

Introduction

Tout d'abord, la langue est un média communicationnel indispensable, que ce soit entre les êtres humains soit entre les humains et les systèmes socio-technologiques intelligents, y compris entre les machines eux-mêmes (Fuller, 2017). L'interaction humaine-machine (IHM) via des interfaces interactives numériques sur Internet est possible grâce à l'affordance de la techno-langue, c'est-à-dire, la société physique est reliée à celle du numérique en reposant non seulement sur les multilingues, mais également sur les multi-technolingués (les langages de programmation) qui tirent majoritairement leurs racines des signes anglo-latins. L'anglais continue probablement à être la langue de mondialisation comblée par Internet depuis des décennies (W3Techs., 2024). Certes, la plupart des inventions numériques sur la Toile depuis 1992 sont dominées par les anglo-saxons. Ainsi, l'expression dynamique Web 3.0 accorde librement aux internautes le contrôle des affaires sociolinguistiques numériques via des agoras de réseaux sociaux centralisés/décentralisés (Facebook, Zoom, Google Meet, YouTube, X, Flickr, LinkedIn, Mastodon, Fediverse, etc.), des moteurs de recherche (Google, Bing, Yahoo, Baidu, Global WP, Exalead, Wikiwix etc.), des applications mobiles (WhatsApp, Instagram, Telegram, Snapchat, etc.), des cryptomonnaies (Bitcoin, Ethereum, Litecoin, Dogecoin, NFTs, etc.) et des métaverses (Second Life, Roblox, Zepeto, Fortnite, etc.). Tous ces éléments renforcent la domination de l'anglais des temps en temps.

Il est clair que l'informatique ne peut rien faire sans ces langages spécialisés et l'anglais, car les usagers fabriquent continuellement des mots en vogue qui favorisent l'anglais dans cette ère des données web, opine Anderson cité dans Hiremath et Kenchakanavar (2016), et l'ère du mobile intelligent où les jeunes se connectent au monde agile. Ces bouleversements soutiennent les IHM, de la réalité virtuelle au monde réel et vice versa, en façonnant les comportements de chacun grâce aux algorithmes (Enjeux RH, 2020).

Cet article se concentre sur les phénomènes techno-linguistiques et multilinguistiques présents sur le weblog du magazine trimestriel français, créé en 2010 par Jérôme Ruskin¹. Le magazine diffuse des informations sur les technologies du futur dans la société réelle. Selon le fondateur, c'est un média qui souhaite faire tester le futur à ses lecteurs. Le magazine tire son nom des personnages curieux des *Lettres persanes* de Montesquieu, parues en 1721², et son but est d'explorer le futur à travers son magazine. Le weblog possède sa propre cryptomonnaie appelée « usbeks (u) »³ pour récompenser les utilisateurs ou les lecteurs du blog. Il s'agit d'une monnaie virtuelle complémentaire avec laquelle les lecteurs peuvent acheter n'importe quoi disponible sur la shop en ligne d'Usbek et Rica. Récemment, Jérôme Ruskin et son équipe éditoriale ont lancé de grandes loteries citoyennes pour expérimenter la réalité du futur. Plus les utilisateurs likent⁴ (aiment), commentent et partagent les articles journaliers, plus ils bénéficient de « usbeks (u) » (voir Figure 5 ci-dessous). En revanche, la condition interdit aux visiteurs de liker, commenter et partager plus de cinq fois le jour⁵. Ces attributs ci-avant appartiennent à l'expression Web 3.0, héritée des fonctionnalités des Web 2.0 et Web 1.0 respectivement.

Cette étude vise également à dévoiler l'homogénéisation de l'anglais comme langue d'Internet, sujet abordé par Barton & Lee (2013) dans leur ouvrage *Language Online: Investigating Digital Texts and Practices*. À travers leurs apports, ils présentent l'anglais comme un amplificateur de la bonne compréhension du monde numérique. Cette domination semble aller à l'encontre de la politique de l'UNESCO, qui encourage l'accès multilingue universel au cyberspace (UNESCO, 2015). Cependant, dans cette concurrence numérique sociolinguistique, le français se présente ouvertement comme un archi-compétiteur de l'anglais en termes démographiques, géographiques, mercatiques et numériques. Malheureusement, l'anglais gagne toujours cette compétition en raison des termes technologiques en vogue provenant de la Silicon Valley aux États-Unis ; des réseaux sociaux et autres consortiums médiatiques numériques anglais qui rendent Internet anglo-centrique. Ce fait est critiqué par Xue et al. (2021), qui ont été inspirés à développer des modèles multilingues mT5 pour renverser la domination de l'anglais dans le prompt engineering, une langue que 80 % de la population mondiale ne comprend pas. En effet, l'anglais occupe 49,7 % du contenu linguistique de tous les sites web, l'espagnol le suit avec 5,9 %, l'allemand 5,4 %, le japonais 4,9 %, et le français est le quatrième avec seulement 4,3 % de l'ensemble du contenu (W3Techs, 2024). Avec tous les pays de la Francophonie et la France elle-même, ils ne peuvent pas générer 10 % du contenu linguistique total sur la Toile. C'est pourquoi, nous étudions également les comportements des utilisateurs vis-à-vis des choix de langage sur Usbek et Rica tout en suivant les répartitions structurelles de cette recherche ainsi : (i) caractérisation des termes, (ii) problématique, (iii) théorie conceptuelle du technolinguisme, (iv) analyse des nuages de mots-clés ou des thèmes d'articles sur Usbek et Rica. Ensuite, (v) études comparatives du multitechnolinguisme et du multilinguisme sur le Web 3.0, et enfin, nous énoncerons les nouvelles perspectives pour susciter le débat sur le multitechnolinguisme.

I. Caractérisation de termes

a. **Technolinguisme** : Ce sont les atouts technologiques et linguistiques de l'impact symbiotique entre langue et technologie (Pfrehm, 2011). Dans la présente recherche, nous le redéfinissons comme des langages de programmation ou de reconnaissance vocale par lesquels l'être humain interagit avec l'informatique. Il ne s'agit pas seulement de la relation mutuelle entre langue et technologie, mais également de l'application linguistique dans l'interface humain-machine (IHM) et l'interface machine-machine (IMM). Nous ajoutons un autre nouveau terme, « multitechnolinguisme » ou « multitechnolinguisme », qui désigne, pour nous, les nombreux langages de programmation utilisés simultanément pour communiquer avec les ordinateurs de manière offusquée.

¹ https://fr.wikipedia.org/wiki/Usbek_%26_Rica

² <https://usbeketrica.com/fr/pages/manifeste> retiré le 17 juillet 2024.

³ <https://usbeketrica.com/fr/pages/manifeste> retiré le 17 juillet 2024.

⁴ Aiment

⁵ <https://usbeketrica.com/shop/>

b. **Multilinguisme** : c'est l'usage simultané de plusieurs langues ; il influence la culture journalière d'une société globale de manière physique ou culturelle numérique.

c. **Web 3.0** : Web 3.0 est inventé par John Markoff de New York Times en 2006 (Hiremath et Kenchakkanavar 2016). La technologie web 3.0 est une Toile sémantique qui a permis l'intégration de données et la personnalisation des comptes sur les sites web d'autrui.

d. **Weblog** : c'est des sites web tricotés des billets publiés à un rythme mélangé (journalier, hebdomadaire et mensuel), intégrant des textes des genres variés, parfois des photos, des podcasts, des vidéos, des agoras et des hyperliens sur des sources d'information externes de nuage de mots-clés.

e. **Webologie** : D'après nous, c'est une étude de technolange, langages naturels, interactions en interface humain-machine ou interface machine-machine sur le site web, weblog et micro-blog dont la « blogologie/blogométrie⁶ » est son sous-genre.

II. Problématique

L'histoire linguistique d'Internet se présente comme celle de la Tour de Babel du Judéo-Christianisme, où, selon la Sainte Bible, il y avait autrefois une seule langue dans le monde (Genèse 11 ; 6-9)⁷. Berners-Lee a inventé une langue destinée au monde virtuel, que la communauté linguistique numérique l'utilise. Bien que l'anglais soit la première langue naturelle à voir le jour sur le site web/blog (Berners-Lee, 1991)⁸, il était programmé en langage de balisage d'hypertexte (HTML). Le site web est bâti sur les réseaux de l'interaction linguistique entre les langages de programmation et les langages naturels, surtout sur l'expression Web 3.0, dont l'humain et l'automate s'en bénéficient en communiquant. Barton et Lee (2013), dans leur ouvrage *Language Online: Investigating Digital Texts and Practices*, soulignent qu'il existe une lacune dans l'étude linguistique numérique qui nécessite une nouvelle enquête (p. i). C'est cette nouvelle piste d'enquête que nous ouvrons ici.

La domination de l'anglais continue dans le contenu linguistique des pages web lisibles en frontend (langue humaine), tandis que le backend reste dominé par les langages de machine. Lee estime que, bien que l'anglais soit une langue répandue sur Internet, il perdait son influence au cours des dernières années (Lee, 2016, p. 118). Cependant, on constate que le contenu linguistique en anglais sur Internet pénètre même les grandes langues du monde telles que le français, l'espagnol, le russe, le chinois, et l'arabe, grâce à l'alternance codique, un style apprécié par les jeunes internautes français sur les réseaux sociaux (La Pige, 2024; Refka, 2021). L'anglais se renforce de plus en plus sur Internet grâce aux réseaux socio-numériques, qui sont le moteur même d'Internet. À cet égard, Pfrehm a baptisé cette situation "technolingualism" (Pfrehm, 2011, p. 11) «technolinguisme en français », car la technologie fait partie de notre vie dans tous ses aspects (Barton & Lee, 2013).

Il est avéré que toute technologie de pointe favorise l'anglais plutôt que le français et d'autres langues. Voilà pourquoi l'alternance de code (anglais-français) sur Usbek et Rica devient le cœur de notre discours numérique. L'alternance de code n'est pas un phénomène nouveau, car de cette alternance de code entre plusieurs langues naît toujours une nouvelle langue. Ainsi, en 1887, Ludwik Lajzer Zamenhof s'est servi de l'alternance de code pour inventer une nouvelle langue, l'espéranto. Cette langue a été développée puis a décliné. Heureusement, grâce à Internet, l'espéranto reprend vie aujourd'hui (Penarredonda, 2018).

⁶ Une étude de blog

⁷ La Traduction œcuménique de la Bible (2000). Genèse 11 :6-9. Stone.

⁸ Le premier site web est lancé en 1991, mais la version définitive est préservée jusqu'à notre jour. (Voir view-source: <https://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html>).

Bien que les experts susmentionnés aient présenté leurs contributions sur ce sujet, nous avons constaté qu'il y a peu d'attention portée sur le langage informatique obscur sur la Toile, qui renforce la domination de l'anglais sur les autres langues mondiales.

III. Cadre conceptuel et méthodologie

Il existe plusieurs théories sur le multilinguisme qui peuvent nous aider à expliquer cette étude. Marx & Hufeisen (2003) proposent trois principales théories du multilinguisme en relation avec la didactique : le modèle de l'alternance linguistique de Björn Hammarberg et Sarah Williams ; le modèle des facteurs de Britta Hufeisen ; et le modèle dynamique du multilinguisme d'Ulrike Jessner et Philip Herdina. Gardner-Cloros (2009) mentionne également trois approches analytiques de l'alternance codique : la description sociolinguistique/ethnographique, l'approche analytique/pragmatique de la conversation, et l'analyse grammaticale des occurrences d'alternance codique. Ces théories se concentrent sur l'apprentissage des langues. Cependant, elles me semblent insuffisantes pour analyser les interactions en alternance de code des internautes, et en particulier, la relation linguistique évolutive qui existe entre le code lisible et illisible des données de contrôle abstrait sur Internet.

Cette recherche s'appuie sur la théorie critique conceptuelle du technolinguisme en utilisant l'approche linguistique écologique de Marie-Anne Paveau (Paveau, 2017). Cette théorie porte sur l'expression multilingue sur Internet, désignée sous le nom de Critique du Code de Site Web, Weblog et Micro-blog (CCSWM), comme une branche de la webologie⁹. Ayant participé virtuellement au Critical Code Studies Working Group 2018 (CCSWG18), une conférence biennale de chercheurs en Critical Code Studies (CCS), Nous avons été inspirés à concevoir une théorie critique conceptuelle en observant un manque dans l'étude critique du code numérique en général. Les experts de CCSWG18 maîtrisaient bien la philosophie de la programmation, mais n'étaient pas spécifiques à un langage informatique particulier. Nous avons proposé lors des débats une critique spécifique du code de site web, qui n'a pas suscité beaucoup d'intérêt parmi les participants. CCSWM est également une critique de l'algorithme du code des sites web ou des blogs. En considérant l'approche linguistique écologique de Paveau, on considère le code source comme un texte digne d'être analysé dans le discours numérique, une idée déjà établie par Sire (2016).

Pour cette étude, comme mentionné ci-avant, nous avons adopté l'approche linguistique écologique, qui examine la relation entre les langages de programmation utilisés pour les sites web, les weblogs et les micro-blogs, et les langues naturelles multilingues sur Internet, en se concentrant sur la production de sens. Les outils d'analyse web et de suivi de trafic tels qu'Alexa, Google Web Analytics, Pyramid Analytics, Yahoo Web Analytics, Real-Time Web Analytics, etc., sont essentiels pour permettre ces études herméneutiques du code web 3.0, au-delà de l'analyse technique des comportements des utilisateurs sur les sites web. Contrairement à une approche purement technique, nous utilisons une critique sémiotique analytique en examinant des captures d'écran d'extraits de code. Cette méthode est appliquée pour analyser la programmation linguistique, l'alternance de code, le discours mixte, ainsi que les interactions des utilisateurs multilingues sur le weblog d'Usbek et Rica.

Les modules Python tels que « Langdetect¹⁰ », « Matplotlib.pyplot¹¹ », et « Pandas¹² » ont été utilisés pour collecter et analyser les données des articles et des thèmes publiés entre le 19 juillet 2016 et le 31 juin 2018, ainsi que du 1er juillet 2018 au 18 juin 2020. À cette époque, les données étaient disponibles sur le sitemap, un outil de développement web qui contient les métadonnées des pages web accessibles au public et aux moteurs de recherche comme Google et d'autres¹³.

La théorie conceptuelle du technolinguisme de CCSWM est pertinente pour cette étude, car Internet relie de plus en plus la société virtuelle à la société réelle, notamment grâce à certains dispositifs

⁹ <http://www.webology.org/about.html>

¹⁰ <https://pypi.org/project/langdetect/>

¹¹ https://matplotlib.org/3.5.3/api/_as_gen/matplotlib.pyplot.html

¹² <https://pandas.pydata.org/>

¹³ <https://developers.google.com/search/docs/crawling-indexing/sitemaps/overview>

numériques, en particulier les smartphones (Miller et al., 2022). De plus, les codes lisibles et illisibles dans les navigateurs web sont également considérés comme des patrimoines culturels (Sire, 2016). L'importance de cette approche, CCSWM, réside dans l'explication des mécanismes linguistiques des interactions dans l'interface homme-machine (IHM) ou l'interface multimodale (IMM) sur Internet. Cette réalité est en constante évolution dans l'étude de la culture numérique.

IV. L'analyse de nuage de mots-clés sur le blog d'Usbek et Rica en 2018

a. Codes de nuage de mots-clés sur Web 3.0

Le web de données (Web 3.0) se caractérise par la possibilité pour les utilisateurs des environnements virtuels de s'engager dans une création collaborative, des débats, des commentaires dynamiques en plusieurs langues, permettant ainsi de corriger et éditer facilement les contenus numériques. C'est un espace où l'on peut partager des idées, aimer une photo ou un post dans une agora virtuelle (Hiremath & Kenchakkanavar, 2016). Les sites web sont développés dans plusieurs langages de programmation nécessaires pour afficher des pages web bien structurées. Le web est avant tout un hôte de « technolangues » multiples (HTML, CSS, JavaScript, etc.), de l'anglais (langue naturelle) et « de toutes les langues » dit Crystal (2006).

Les balises du web sont souvent encapsulées dans des éléments « <div> </div> ». Le nuage de mots-clés, ou « tag cloud », est utilisé dans les blogs pour indiquer les principaux sujets ou les hyperdocuments publiés par un magazine. Il aide les utilisateurs à choisir leurs préférences pour une lecture ultérieure ou à relier des sujets similaires dans un cadre cohérent. Certains chercheurs définissent le nuage de mots-clés comme « une représentation, dans une liste alphabétique, des mots-clés les plus fréquents, avec une taille de caractère proportionnelle à leur utilisation » (Hardy & Palermo, 2010, p. 79). Le code ci-dessous affiche les mots-clés hiérarchiquement alphabétisés. Les blogueurs d'Usbek et Rica ont adopté cette architecture pour leur weblog. De plus, les symboles de notation tels que « / », « [] », « { } », « < > », la balise ouvrante « < » et la balise fermante « > », la balise HTML « /tag/ » ou le signe dièse « # », etc., sont des éléments de notation symbolique ancrés dans le langage de programmation anglais, qui masquent les codes sources tout en affichant ceux qui sont lisibles sur une page web.

Aujourd'hui, grâce au Web 3.0, la programmation intelligente permet la reconnaissance faciale des utilisateurs dans les photos et vidéos numériques, ainsi que le partage des données privées avec d'autres sites web à des fins publicitaires selon les besoins des visiteurs (Carter, 2018). Par conséquent, les technologies évoluent vers un langage agile.

La page web illisible est construite dans une IMM multimodale dont le langage est technolange, tandis que la page web lisible est construite dans une IHM. C'est ainsi que d'une part, les humains communiquent avec l'informatique sans difficulté, et d'autre part, les systèmes informatiques communiquent entre eux. Le langage de haut niveau, utilisant la syntaxe, la sémantique et la morphologie reconnues, affiche des langages humains lisibles basés sur des mots-clés simplifiés de l'anglais, c'est-à-dire des mots abrégés qui paraissent codiques en nomenclature. Cette nomenclature renforce la prédominance de l'anglais dans un contexte multilingue/multitechnologique, manifestant ainsi une forme de domination linguistique postcoloniale par rapport à d'autres langues mondiales (Marino, 2018). À ce sujet, Berry (2008) insiste à dire que:

The *delegated code* is written by the programmer as these highly abstracted languages use a formalised syntax and are usually constructed around simplified English keywords. Together with symbols and punctuation, programs are written in a structured syntactical style made up of statements, loops and conditionals by the programmer to construct the logical operation of the program. (Berry, 2008)¹⁴

C'est à ce stade que le code délégué est écrit par le programmeur, car ces langages hautement abstraits utilisent une syntaxe formalisée et sont généralement construits autour de mots-clés anglais

¹⁴ <https://thirteen.fibreculturejournal.org/fcj-086-a-contribution-towards-a-grammar-of-code/>

simplifiés. Avec les symboles de notation et la ponctuation, les programmes sont écrits dans un style syntaxique structuré composé d'instructions, de boucles et de conditionnels par le programmeur pour construire l'opération logique du programme (notre traduction).

Le site web représente un univers linguistique à la fois pour l'homme et pour la machine. Par exemple, le code de programmation ci-dessous montre un extrait du code source de web illustrant l'alternance codique entre le français et l'anglais, tiré du site usbeketrica.com. Les codes lisibles produisent des structures descriptives et logiques.

```
<div>
  <ul id="tags">
    1. <li><a href="/tag/animal">Animal</a></li>
    2. <li><a href="/tag/art">Art</a></li>
    3. <li><a href="/tag/climat">Climat</a></li>
    4. <li><a href="/tag/corps">Corps </a></li>
    5. <li><a href="/tag/culture">Culture </a></li>
    6. <li><a href="/tag/democratie">Démocratie </a></li>
    7. <li><a href="/tag/design">Design </a></li>
    8. <li><a href="/tag/economie">Économie </a></li>
    9. <li><a href="/tag/education">Éducation </a></li>
    10. <li><a href="/tag/energie">Énergie </a></li>
    11. <li><a href="/tag/environnement">Environnement </a></li>
    12. <li><a href="/tag/espace">Espace </a></li>
    13. <li><a href="/tag/food">Food </a></li>
    14. <li><a href="/tag/geopolitique">Géopolitique </a></li>
    15. <li><a href="/tag/hacker">Hacker </a></li>
    16. <li><a href="/tag/histoire">Histoire </a></li>
    17. <li><a href="/tag/innovation">Innovation </a></li>
    18. <li><a href="/tag/internet">Internet </a></li>
    19. <li><a href="/tag/jeux-vidéo">Jeux vidéo </a></li>
    20. <li><a href="/tag/maker">Maker </a></li>
    21. <li><a href="/tag/media">Média </a></li>
    22. <li><a href="/tag/mobilite">Mobilité </a></li>
    23. <li><a href="/tag/mode">Mode </a></li>
    24. <li><a href="/tag/numerique">Numérique </a></li>
    25. <li><a href="/tag/podcast">Podcast </a></li>
    26. <li><a href="/tag/politique">Politique </a></li>
    27. <li><a href="/tag/religion">Religion </a></li>
    28. <li><a href="/tag/robot">Robot </a></li>
    29. <li><a href="/tag/sante">Santé </a></li>
    30. <li><a href="/tag/science">Science </a></li>
    31. <li><a href="/tag/science-fiction">Science-Fiction </a></li>
    32. <li><a href="/tag/societe">Société </a></li>
    33. <li><a href="/tag/sport">Sport </a></li>
    34. <li><a href="/tag/technologie">Technologie </a></li>
    35. <li><a href="/tag/transhumanisme">Transhumanisme </a></li>
    36. <li><a href="/tag/travail">Travail </a></li>
    37. <li><a href="/tag/utopie">Utopie </a></li>
    38. <li><a href="/tag/ville">Ville </a></li>
  </ul>
</div>
```

Figure 1: Extrait du code source web d'Usbek et Rica capturé le 19 mars 2018.

Dans cet extrait de code, les mots en gras représentent le français, les mots argentés indiquent des éléments de code spécifiques à l'anglais, et ceux en gris (tels que design, food, hacker, Internet, marker, podcast, et science-fiction, et sport) sont des termes anglais empruntés aux français via des codes mixtes ou l'alternance codique affichés sur la page web lisible. Les mots français entre balises peuvent être remplacés par n'importe quelle autre langue, telle que le russe, le portugais, le chinois, le margi, l'igbo, le yorouba, l'haoussa, etc., à condition que les graphèmes, glyphes, logogrammes ou symbogrammes de la langue respectent les standards du système d'orthographe numérique, et soient encadrés par des guillemets simples ou doubles anglais « ' ou " ».

Deux ou plusieurs thèmes de nuage de mots-clés sont reliés par une balise d'ancrage HTML « <a> » (anchor tag) utilisant l'attribut « href », qui sert de référence de lien « head reference » (Voir Figure 1.0), et se terminent par une balise fermante « ». Le nuage de mots-clés sur le site d'Usbek et Rica a été archivé pour la période du 19 juillet 2016 au 31 mars 2018. Durant cette période, 648 508 utilisateurs ont interagi avec les articles publiés. Parmi eux, 4 327 utilisateurs ont partagé leurs articles préférés avec des amis, et 155 954 articles ont été marqués pour une lecture ultérieure. Ces interactions, comme les mentions "j'aime", le partager et la lecture ultérieure, sont enregistrées sous la forme d'un « graphopticon » sur le site d'Usbek et Rica. Le graphopticon montre les articles les plus lus et interactifs chez les lecteurs (voir Figure 2).



Figure 2: Graphopticon sur Usbek et Rica. Prise d'écran le 29 mars 2018.

Les commandes placées entre crochets, la classe « string » en programmation, et le balisage « </tag> » encadré par des guillemets anglais constituent les points de connexion entre les multitechnolangues et les multilingues. Ces éléments représentent des classes symboliques en programmation, qui instruisent l'ordinateur à afficher les mots entre guillemets anglais de manière lisible, tels qu'ils sont, sans altération, comme indiqué par Waliya (2023). Sur le web, le code existe en deux dimensions : le code statique, lisible pour les humains, et le code fonctionnel, destiné aux machines (Berry, 2008). En d'autres termes, les automates lisent et comptent les utilisateurs/visiteurs, les mentions "j'aime", les partages, les commentaires et les messages envoyés par courriel. Ils alertent les destinataires du courriel en signalant qu'un message a été reçu/ajouté dans leur boîte de réception. Ces interactions dans l'IMM sont rendues possibles grâce au compilateur, qui traduit les langages naturels en langage binaire (0 et 1) pour les machines. Ces métadonnées sont configurées dans les dispositifs et les codes de base opine Marino (2018). L'alphabet latin et les mots anglais (comme class, load, tag) restent omniprésents dans les langages de programmation, même en arabe, en chinois, etc. Par exemple, les extraits de code des Figures 3 et 4 montrent des mots entre balise ouvrante et balise fermante de division « div class= "views-field-field-abstract-value"...</div> » sont en anglais. Ces mots constituent les codes de base pour les weblogs populaires en arabe et en chinois, dont les orthographes sont logogrammes, illustrant ainsi la domination hégémonique de l'anglais sur Internet et dans la programmation des weblogs.

```
<div class="views-field-field-abstract-value">
  <div class="field-content">التفوية الامباراه فدي الاعدي الاتقص رخم ل هدف ب هدفيد فون ابد ن طيف ب هدف تناخر ه الامانيا مبخو ولي
  16. الا دور بلوه فدي املا ه التلقو ح املا ل بحيد 2018 الا علامه كفاك مد اتسادسه الامجمو عه ل حساب لا اولميد فديشه ملجج جرد التو
  </div>
</div>
```

Figure 3: Extrait du code source web d'albawaba¹⁵ le 24 juin 2018

```
<map name="Map1" id="">
  <area shape="rect" coords="173, 81,299,137" href="http://home.sina.com" target="_self" alt="北美新浪" title="北美新浪" onclick="_S_uaTrack ('global guide', 'us');" />
  <area shape="rect" coords="468, 81,572,129" href="http://www.sina.com.cn" target="_self" alt="北京新浪" title="北京新浪" onclick="_S_uaTrack ('global_guide', 'beijing');" />
</map>
```

Figure 4: Extrait du code source web de sina¹⁶ le 24 juin 2018

¹⁵ <https://www.albawaba.com>

¹⁶ <https://www.sina.com.cn>

Les extraits de code ci-dessus révèlent que l'anglais constitue la base linguistique de la programmation, que la machine comprend et suit via des algorithmes grammaticaux. Cela explique comment l'anglais a envahi les pages web affichées ainsi que le code source du web dans son intégralité. Le nuage de 38 thèmes contient 8 thèmes anglais (Figure 5):



Figure 5 : Le nuage de mots-clés des 38 thèmes d'articles sur Usbek et Rica depuis 2016-2018. (Prise d'écran du navigateur Chrome- le 19 mars 2018)

Par la suite, le nombre de thèmes est passé de 38 à 77, une augmentation de plus de 100% entre 2020 et 2024 faisant 13 thèmes anglais (Figure 6).



Figure 6 : Le nuage de mots-clés des 77 thèmes d'articles sur Usbek et Rica (Prise d'écran du navigateur Chrome- le 17 juillet 2024)

L'intégration des données du web sémantique est interprétée différemment par les technologies des navigateurs web. Jusqu'à présent, l'expression « web 3.0 » n'a pas réussi à résoudre le problème de déchiffrement uniforme dans les navigateurs web (voir les captures d'écran de Chrome et Firefox, Figures 5 et 6). Ces navigateurs, qui agissent comme des robots de lecture pour les sites web, les weblogs et les micro-blogs, n'interprètent pas les pages d'Usbek et Rica de manière identique. Ce phénomène, caractéristique des textes numériques, est appelé « labilité » par Bootz (2022, p. 75), car la labilité dynamise le sens des signes affichés à l'écran. Par exemple, sur Firefox, le code de cadre avec #Ville (voir Figure 6) apparaît au cinquième rang avec un autre nuage de mots-clés qui se termine horizontalement à droite. En revanche, sur Chrome, #ville apparaît en bas, à la fin d'une colonne verticale distincte à gauche (Figure 5.0). Cela montre que les deux navigateurs web, au fil des années, affichent le cadre « <div> </div> » différemment, indiquant que les systèmes informatiques ne présentent pas une page web lisible de manière uniforme.

b. Le nuage de mots-clés en diagramme de dispersion sur web 3.0

La typologie des mots-clés sur Usbek et Rica reflète un maillage dynamique. Le web 3.0 organise les thèmes des articles quotidiens, hebdomadaires et mensuels en ordre alphabétique, en mettant en avant les mots-dièses (hashtags) marqués par le signe typographique croisillon « # ». Ces hashtags servent d'indicateurs visuels pour attirer l'attention des visiteurs du blog d'un simple coup d'œil. Parmi les mots-clés dièses associés aux thèmes (Figure 5.0), on observe une présence notable de mots anglais en vogue, ou « buzzwords », tels *qu'hacker, sport, maker, podcast, design, food, science-fiction et Internet* (Figure 7) qui pénètrent les thèmes français.

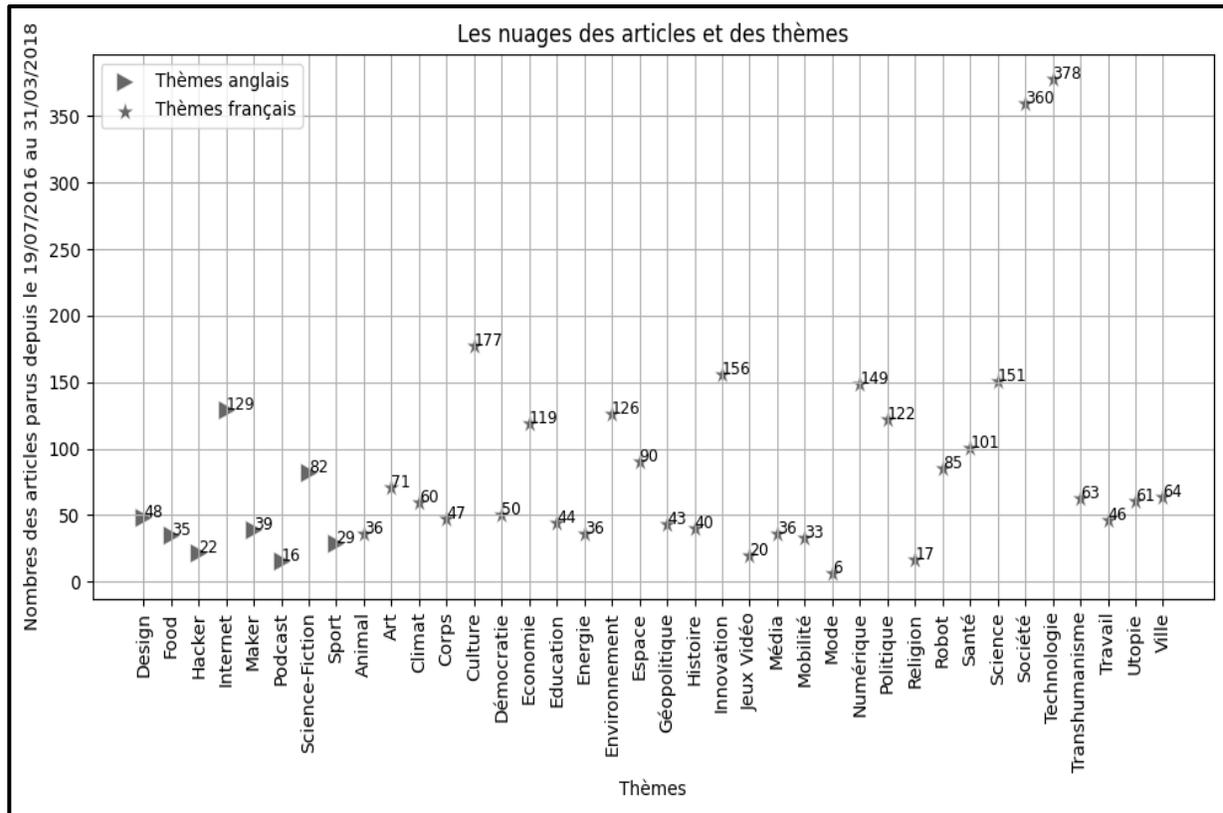


Figure 7 : le nuage de mots-clés des thèmes sur Usbek et Rica entre 19 juillet 2016 et 31 mars 2018.

Les points triangulaires et astérisques de Figure 7 illustrent la répartition des 3 187 articles divers (textes, vidéos et photos) publiés sur Usbek et Rica entre le 19 juillet 2016 et le 31 mars 2018. Ces articles sont interconnectés, partageant les mêmes titres reliés à deux ou trois thèmes du nuage de mots-clés (Figure 2). Les nœuds triangulaires représentent les thèmes en anglais, tandis que les nœuds astérisques représentent les thèmes en français. Le numéro attaché à chaque point ou nœud indique la somme totale d'articles associés à chaque thème. Entre les années mentionnées, les thèmes en anglais ont connu une augmentation de 21,05% vis-à-vis de 78,95 % thèmes en français sur Usbek et Rica. En 2018, les 38 thèmes sont liés à 2 990 textes, 165 galeries de photos et 32 vidéos via des hyperliens. Comme indiqué sur le diagramme de dispersion ci-avant, la somme d'articles de thèmes de technologie est le plus élevé avec la valeur de 378 articles contre le plus bas valeur 6 articles de thèmes de la mode. Autrement dit, au cours des années le magazine n'est pas changé son focus mais projeté l'alternance code. Par exemple, les articles enregistrés depuis sa présence sur Internet jusqu'au 31 mars 2018 révèlent certaines alternances codiques sur le blog d'Usbek et Rica, exemplifiant une présence significative de termes anglais. Il faut noter que les mots ou phrases anglaises sont rendus en gros comme le suit :

- i. Pour une gynécologie *do it yourself* (publié le 19 aout, 2016 sur <https://usbeketrica.com>).
- ii. *United colors of Superhéros* (publié le 24 aout, 2016 sur <https://usbeketrica.com>).

- iii. Reportage au *Web Summit* de Lisbonne, entre grisaille et Singularité (publié le 15 novembre 2016 sur <https://usbeketrica.com>).
- iv. *Bye bye* vérité, bonjour mauvaise foi (publié le 18 janvier 2017 sur <https://usbeketrica.com>)
- v. 2035 ou le *very bad trip* des services de renseignement américains (publié le 24 janvier, 2017 sur <https://usbeketrica.com>).
- vi. *Crowd*, le *podcast* : « On consomme la musique plus qu'on ne l'écoute » (publié le 15 décembre, 2017 sur <https://usbeketrica.com>).
- vii. *Fake news*, harcèlement : Terra Nova propose un « service public de la notification » (Publié le 24 janvier, 2018 sur <https://usbeketrica.com>).
- viii. Cellarius, la première œuvre *cyberpunk* 100% collaborative sur *blockchain* (publié le 7 mars 2018 sur <https://usbeketrica.com>).
- ix. *Yellowstone*, avant et après le réchauffement climatique (publié le 12 mars, 2018 sur <https://usbeketrica.com>) etc.

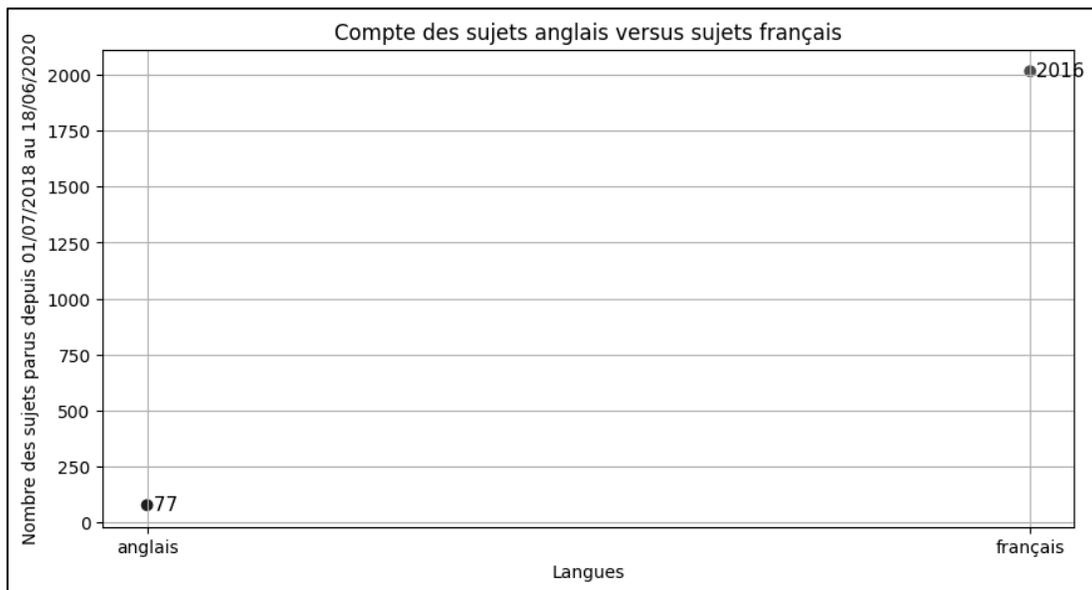


Figure8

Figure 8 : Comptes des articles parus entre 2018 et 2020

Depuis 2018 jusqu'à 2024, 39 thèmes ont été ajoutés, parmi lesquels 5 mots sont en anglais—*blockchain*, *data*, *VR*, *Web*, et *leadership* (voir Figure 6.0). Quant aux articles présentant une alternance codique (français-anglais), il y en a 400 sur les 3 187 publiés entre le 19 juillet 2016 et le 31 mars 2018. L'alternance codique représente 0,13% des articles publiés sur le blog trimestriel parisien. Il est évident que l'anglais occupe 49,7% des contenus linguistiques de tous les sites web à l'échelle mondiale, comme le rapporte la statistique de W3Techs (2024). Nous avons constaté que les données statistiques de l'analyse montrent clairement que l'homogénéisation de l'anglais en tant que langue du web est évidente. L'anglais semble déjà se diriger vers le statut de lingua franca de la mondialisation. Comme souligné un homme politique français, « ...l'anglais est une clé d'accès à la technique et à la modernité, ce qui est une évidence ». (Gallo, 2007, p. 63). La modernité, dans ce contexte, se réfère à Internet, car « ...le monde n'est plus un village, il est devenu un réseau. La concurrence y est féroce et la réussite dépend d'une parfaite maîtrise de l'information, ainsi que de l'audace et de la rapidité. » (Josselin, 2007, p. 144). Au fil du temps, Usbek et Rica a subi une augmentation de l'alternance codique de 0,037 % (77 articles anglais sur 2016 articles français) entre le 1 juillet 2018 et le 18 juin 2020 (voir Figure 8). Bien que ce chiffre puisse sembler insignifiant, il indique une prééminence du vocabulaire anglais sur le français, notamment en ce qui concerne les termes technologiques présents sur le blog.

V. Études comparatives de technolinguisme et multilinguisme sur le web 3.0.

Les technolangues populaires qui configurent le Web incluent : les langages de balisage d'hypertexte (HTML versions 1 à 5) / Markdown, XML (Extensible Markup Language), Perl, les feuilles de styles en cascade (CSS3), JavaScript (JS), Java, PHP (Hypertext Preprocessor), SQL (Structured Query Language), Python, C++, C, C#, Objective-C, ASP.NET, fichiers statiques, Silverlight, ColdFusion, Go, Ruby, etc. Ces langages de programmation se comparent à des langues naturelles telles que l'anglais, le russe, l'allemand, le japonais, l'espagnol, le français, etc. Ils sont intégrés pour produire une Interface Homme-Machine (IHM).

Mon argument est que les langages de programmation peuvent s'intégrer dans un multilinguisme cybernétique, également appelé multitechnolinguisme, où les machines interprètent et communiquent les codes entre elles pour créer une page web bien structurée et dynamique. Les langages de programmation web fonctionnent comme des langues naturelles. En effet, « comme tout langage, HTML est soumis à une syntaxe spécifique, qui définit l'ordre dans lequel les balises doivent ou peuvent apparaître dans une page web ». (Ed Tittel, 2013).

Donc, l'objectif de comparer le multilinguisme et le multitechnolinguisme dans le contexte du Web 3.0 est admissible. Les deux concepts nécessitent des lexèmes, une grammaire structurée, et des signes de ponctuation. Le multilinguisme concerne la connexion entre plusieurs langues naturelles, tandis que le multitechnolinguisme se concentre sur les langages informatiques. Bien que les langages de programmation soient conçus selon un modèle linguistique humain, ils ne fonctionnent pas comme des langues naturelles.

En comparaison, le multilinguisme et le multitechnolinguisme ne partagent pas tous les modes verbaux. Les modes qui préoccupent la technolange sont principalement les modes impératifs et conditionnels, utilisés pour donner des instructions aux machines ou pour contrôler l'abstraction algorithmique des données.

De plus, la syntaxe des technolangues est généralement beaucoup plus concise que celle des langages naturels (Saujani, 2017). En revanche, les langues naturelles comme le français et l'anglais utilisent une gamme complète de modes verbaux et des règles grammaticales plus complexes que celles des langages de programmation.

À l'inverse, les utilisateurs monolingues seront de plus en plus assistés à l'avenir grâce à des logiciels et applications intelligents tels que ChatGPT, Gemini, Skype Translator, Google Translator, Poliglu et autres outils de traduction automatique intégrés au Web 3.0. Ces outils sont conçus pour fournir des traductions instantanées dans toutes les langues disponibles sur Internet, y compris les sites web, weblogs, et micro-blogs. C'est pourquoi Rivierre (2018)¹⁷ souligne que « le plaisir d'apprendre une autre langue aura totalement disparu, avec lui la possibilité de faire travailler notre mémoire, de se confronter à d'autres cultures, et de voir le monde autrement ». Ainsi, l'informatique pourrait remplacer certaines fonctions humaines, et les langues en concurrence avec l'anglais pourraient perdre leur force concurrentielle, car la Silicon Valley continue de surveiller et de dominer le développement technologique.

Pour résoudre le problème de concurrence entre les langues mondiales sur le net, il est crucial que chaque langue développe son propre langage de programmation, sinon l'anglais risque de dominer toutes les autres (Leppänen & Saija, 2012). Une solution envisageable est l'invention d'un langage de programmation post-symbolique pour la réalité virtuelle, comme celui proposé par Jaron Lanier¹⁸. Ce langage n'utilise pas de symboles de notation, de numéros, de ponctuation ou de lettres ; au lieu de cela, il se base sur le dessin comme méthode de programmation (vradmin, 2017)¹⁹. Le dessin, étant neutre et universel, pourrait potentiellement remplacer les langages textuels.

¹⁷ Adrien Rivierre (03/01/2018) <https://usbeketrica.com/article/tous-polyglottes-avec-la-realite-virtuelle> consulté le 14/02/18.

¹⁸ Virtual Reality Society. <https://www.vrs.org.uk/virtual-reality-applications/programming-languages.html>. consulté le 28 juin 2018

¹⁹ *ibid*

Conclusion

Enfin, notre analyse a présenté un panorama des concurrences entre l'anglais et d'autres langues, en particulier le français, son analogue européen, dans le code source jusqu'à l'affichage transitoire observable à l'écran du blog Usbek et Rica. L'invasion de l'anglais constitue une menace pour la langue française, ainsi que pour d'autres langues, car la majorité des machines ne compilent que des codes basés sur des langages de programmation ou des technolectes en anglais, comme l'ont montré les exemples précédents.

Les codes de programmation fondamentaux et les termes technologiques populaires renforcent cette situation, car les internautes issus de diverses cultures préfèrent utiliser l'anglais pur ou l'alternance de code sur le web, dans les blogs, les forums, les interfaces de commentaires et autres fonctionnalités.

Il est évident que, malgré la diversité géographique des interlocuteurs en ligne, l'anglais reste la langue de prédilection des internautes à travers le monde. En effet, une sorte de lingua franca planétaire semble se dessiner. Ainsi, l'alternance de code sur le blog d'Usbek et Rica entraîne une augmentation exponentielle des interactions entre blogueurs, facilitée par des interfaces interactives qui exploitent l'intelligence artificielle, notamment l'IA générative.

Il est donc avéré que le technolinguisme pourrait, à terme, supplanter les langues quotidiennes, car l'évolution se dirige déjà vers l'usage de langages numériques textuels, imaginaires, sonores, sémiotiques (comme les emojis et les émoticônes) et vers le « graphopticon », grâce à une génération de natifs numériques qui vivent pleinement en ligne.

Peut-on envisager que l'anglais devienne obsolète sur internet dans un avenir proche ? Quelle langue sera utilisée pour programmer les automates dans le futur Web 4.0 ? Ce dernier sera-t-il fondé sur le Web 3.0, via des interfaces de réalité virtuelle ou augmentée capables de lire les intentions humaines à partir des impulsions cérébrales, comme le postulent Keshab et Raja (2015) ? Cela semble possible uniquement dans les métavers, cet internet du futur où les barrières linguistiques pourraient être effacées. Il est donc essentiel que les politiques linguistiques veillent à préserver les langues naturelles, qui sont une part fondamentale de notre identité culturelle.

BIBLIOGRAPHIE :

- Barton, D., & Lee, C. (2013)** *Language Online Investigating Digital Texts and Practices*. Routledge
- Berry, D. M. (2008)** A Contribution Towards a Grammar of Code. // *The Fibreculture Journal*, 86 (13). <http://thirteen.fibreculturejournal.org/fcj-086-a-contribution-towards-a-grammar-of-code>
- Bootz, P. (2022)** Procedural Model Applied to the Analysis of Programmed Digital Literary Works. In M. El Fali (Éd.), *Exploring Contemporary Digital Poetics* (p. 65-97). Laboratoire de Langue, Littérature, Imaginaire et Esthétique.
- Carter, D. (2018)** Surf the web in privacy. In D. Carter (Éd.), *Stay Safe and Private Online*. Old Town Publishing
- Crystal, D. (2006)** *Language and the Internet* (2^e éd.). University of Cambridge.
- Danisi, M. (2017)** *The Semiotics of Emoji: The Rise of Visual Language in the Age of The Internet*. Bloomsbury Academic.
- Ed Tittel, C. M. (2013)** *Programmation HTML5 avec CSS3 pour Les Nuls* (D. Denis & B. Stéphane., Trad.). Wiley Publishing.
- Enjeux RH. (2020, août 20)** *Mon nouveau collègue l'algorithme*. Enjeux RH. <https://www.enjeuxrh.com/mon-nouveau-college-lalgorithme>
- Fuller, M. (2017)** *How to Be a Geek: Essays on The Culture Of Software* (1er éd.). Polity Press.
- Gallo, M. (2007)** L'université française face à la construction de l'Europe. In J. Girardet, *Panorama 4 : Méthode de Français*. (P. 63). Clé Internationale.
- Gardner-Cloros, P. (2009)** *Code-switching*. University of Cambridge Press.
- Hardy, J.-M., & Palermo, G. (2010)** *Réussir son web en 60 fiches* (3^e éd.). Dunod.

- Hiremath, B. K., & Kenchakkanavar, A. Y. (2016)** An Alteration of the Web 1.0, Web 2.0 and Web 3.0: A Comparative Study. // *Imperial Journal of Interdisciplinary Research (IJIR)*, 2(4), 705-709.
- Josselin, C. (2007)** L'ambition francophone. In J. Girardet, *Panorama 4 : Méthode de Français*. (p. 144). Clé Internationale.
- Keshab, N., & Raja, I. (2015)** What Comes After Web 3.0? Web 4.0. In J. K. Saha, G. Kandar, D. Maji, & A. K. Mandal (Éds.), *Computing and Communication Systems*.
- La Pige. (2024, février 14)** *Abus d'anglicismes chez les jeunes : En français, please!* Le Quotidien. <https://www.lequotidien.com/partenaires/la-pige/2024/02/14/abus-danglicismes-chez-les-jeunes-en-francais-please-VFTKK2NKURCO3IEJ3WZ5IQAYYM/>
- Lee, C. (2016)** Multilingual resources and practices in digital communication. In A. Georgakopoulou & T. Spilioti (Éds.), *The Routledge Handbook of Language and Digital Communication* (pp. 118-121). Routledge Taylor & Francis.
- Leppänen, S., & Saija, P. (2012)** Multilingualism on the Internet. In M. Martin-Jones, A. Blackledge, & A. Creese (Éds.), *The Routledge Handbook of Multilingualism* (p. 384-402). Routledge Taylor & Francis.
- Marino, M. C. (2018)** Culture through Code. *Routledge*, 475.
- Marx, N., & Hufeisen, B. (2003)** Multilingualism: Theory, Research, and Didactics. In B. Gerd & K. Sanders (Éds.), *New Visions in Foreign and Second Language Education* (p. 181-192). LARC Press.
- Miller, D., Rabho, L., Awondo, P., de Vries, M., Duque, M., Garvey, P., Haapio-Kirk, Hawkins, C., Otaegui, A., Walton, S., & Wang, X. (2022)** *Le Smartphone Global: Au-delà d'une culture jeune - A French Translation of the Global Smartphone* (S. Engola, Trad.). UCL Press. <https://doi.org/10.14324/111.9781800081505>
- Paveau, M.-A. (2017)** *L'analyse du discours numérique : Dictionnaire des formes et des pratiques*. Hermann Éditeurs.
- Penarredonda. (2018)** *The invented language that found a second life online* [Web]. <http://www.bbc.com/future/story/20180110-the-invented-language-that-found-a-second-life-online>
- Pfrehm, J. (2011)** *Technolingualism: The Mind and the Machine*. Bloomsbury Academic.
- Refka, B. A. (2021)** L'emploi des anglicismes par les internautes francophones : Cas des réseaux sociaux. *Akofena*, 1(002), 199-206.
- Saujani, R. (2017)** *Girls Who Code : Learn to Code and Change the World*. Viking.
- Sire, G. (2016)** Ce que coder veut dire : Y a-t-il un langage de programmation ? Plaidoyer pour une approche sociolinguistique du code. *French Journal For Media Research*, 6, 1-6.
- UNESCO (2015)** Multilingualism in Cyberspace: Indigenous Languages for Empowerment. *UNESCO Regional Conference For Central America*. Multilingualism in Cyberspace, San José.
- vradmin (2017, mai 5)** Virtual Reality Programming Languages. *Virtual Reality Society*. <https://www.vrs.org.uk/virtual-reality-applications/programming-languages.html>
- W3Techs (2024, juillet 15)** *Usage Statistics and Market Share of Content Languages for Websites* [Blog]. W3Techs: Web Technologies Survey. https://w3techs.com/technologies/overview/content_language
- Waliya, Y. J. (2023)** *Twittérature : Analyse Technodiscursive de la Twitterbot Poésie de Leonardo Flores* [Thèse de Master]. Ahmadu Bello University.
- Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A., & Raffel, C. (2021)** *mT5: A massively multilingual pre-trained text-to-text transformer* (arXiv:2010.11934). arXiv. <https://doi.org/10.48550/arXiv.2010.11934>