

« Avec le numérique, l'hégémonie de l'anglais devient de l'histoire ancienne. »

L'anglais va certainement jouer un rôle important dans le façonnage du nouvel ordre linguistique mondial mais son impact majeur sera de donner naissance à de nouvelles générations de locuteurs bilingues et multilingues partout dans le monde.

David Graddol, « The Future of Language », *Science*, 2004

Si la langue est le véhicule privilégié de la culture, qu'en est-il dans le monde numérique et dans la société ?

L'influence de l'anglais a cru au cours du XX^e siècle du fait de la domination économique et politique des États-Unis. Il est devenu la principale langue de communication internationale et contribue à l'instauration d'une forme d'hégémonie culturelle étroitement associée à l'économie libérale et à la globalisation des échanges marchands. Son emploi est parfois considéré comme idéologique ; il marque la modernité et l'ouverture internationale. Il est aussi contesté : résistance à l'hégémonie culturelle par la défense des langues nationales ou du plurilinguisme ; défense d'une langue internationale plus accessible comme l'esperanto ; équité politique (travailler en anglais revient à donner un avantage compétitif à ceux dont c'est la langue maternelle).

Dans ce contexte, Internet est-il le vecteur d'une domination universelle de l'anglais ou favorise-t-il la pluralité linguistique, facilitant le passage d'une langue à l'autre grâce aux outils de traduction automatique et préservant les langues menacées de disparition. À ses débuts, en 1969, il est

exclusivement anglophone, mis en place par le Pentagone puis dans les organismes gouvernementaux et les universités américaines. Après 1989, avec la création du web, puis des moteurs de recherche, il se développe d'abord aux États-Unis puis dans le reste du monde. L'anglais y domine alors que seulement 9 % de la population mondiale est anglophone de naissance. Le numérique ne reflète pas la diversité linguistique comme on le constate aussi avec le projet Gutenberg qui, en 2011, avait numérisé et mis gratuitement à disposition près de 31 000 livres en anglais, pour seulement 1 800 en français, 800 en allemand, 600 en finois, 500 en néerlandais et portugais, 400 en chinois, 300 en espagnol, etc.

La diversité linguistique s'accroît cependant. En 2015, si l'anglais domine encore avec 850 millions d'internautes, son poids relatif s'érode passant de 75 % en 1998 à 45 % en 2008, puis 26 % en 2015. Il est suivi de près par le chinois (21,5 %), en forte croissance. Puis viennent l'espagnol (7,5 %), l'arabe (4,8 %), le portugais (4 %), le japonais (3,5 %), le russe (3,2 %), le malais (2,9 %), le français (2,8 %) et l'allemand (2,6 %). Cette diversité linguistique reste toutefois relative ; ces dix langues représentent environ 80 % des internautes et des contenus accessibles sur le web alors que 6 000 langues sont parlées dans le monde.

La décroissance de l'anglais et la montée en puissance d'autres langues laissent penser que le numérique favorise toutefois le multilinguisme. Les grands acteurs du web y contribuent en donnant effectivement plus de poids aux langues locales ; le moteur de recherche de Google en supporte 345 depuis 2009, Wikipédia 285, Facebook 70 en 2012, Twitter 21 et LinkedIn 17. Par ailleurs, les humanités

numériques et l'UNESCO avec son projet Babel favorisent la conservation de langues menacées de disparition. Pour l'UNESCO, il s'agit de défendre la diversité linguistique et culturelle, notamment en développant des outils multilingues en ligne et les outils de traduction automatique, en adaptant les systèmes d'indexation multilingues des sites web et outils terminologiques, en renforçant l'interopérabilité, en favorisant la numérisation et la mise en ligne des archives et en agissant sur les politiques nationales et internationales.

Au sein des humanités numériques, malgré leur vocation à travailler sur les patrimoines culturels de l'humanité, la situation semble aussi dominée par le monde anglo-saxon, favorisant, de fait, l'hégémonie de l'anglais. Les premiers centres de recherche surgissent aux États-Unis, principalement sur la côte Est. Le premier, créé en 1994, à l'Université de Virginie entre les départements d'histoire et d'informatique, et la bibliothèque, pour accueillir plusieurs projets qui marquent l'histoire des DH : *The Valley of the Shadow*, *l'Archive Rossetti*, *l'Archive William Blake*, le développement du langage d'indexation TEI et d'outils de visualisation et d'annotation. D'autres centres de ce type sont ensuite créés à l'Université du Maryland, à UCLA (à l'origine du premier manifeste pour les DH), à l'Université du Michigan, à Georges Mason (conception des outils en accès libre Zotero pour la gestion de références bibliographiques et Omeka pour la gestion de bases de données, création des premières non conférences THATCamp), etc. Ils se développent aussi ailleurs dans le monde anglophone au Canada (Victoria, McGill à Montréal), en Australie (Melbourne, Newcastle), en Irlande (University College à Dublin) et en Grande-

Bretagne (Oxford, Glasgow, UCL et King's College à Londres, Birmingham). Les pays non anglophones, bien qu'ils soient engagés dans de vastes projets de numérisation, sont initialement marginaux en Allemagne (Cologne, Gottingen), dans les pays scandinaves, en Suisse (Bern, Zurich, Lausanne, et le portail des sciences historiques de l'Académie suisse des sciences humaines et sociales Infoclio), puis dans les pays latins (Espagne, Italie, France).

Les sociétés savantes reflètent aussi l'hégémonie anglo-saxonne. La revue *Literary and Linguistic Computing* est créée en Grande-Bretagne en 1986. L'association correspondante (ALLC) devient, en 2012, The European Association for Digital Humanities (EADH) mais reste dominée par les Britanniques. En 1978, est née l'Association for Computers and the Humanities (ACH), basée aux États-Unis, pour animer les échanges notamment avec la foire aux questions *DH Answers* et la revue *Digital Humanities Quarterly*. Les différentes associations se fédèrent progressivement depuis 2002 pour former l'Alliance of Digital Humanities Organizations (ADHO) qui organise la conférence annuelle *Digital Humanities DH*. Sur le continent, les chercheurs italiens, espagnols et allemands se structurent progressivement et fondent leurs associations nationales en 2011 et 2012. La communauté latino-américaine se structure également dans la même période. En 2014, est créée l'association francophone Humanistica. L'EADH lance par ailleurs le groupe Global Outlook (GO : DH) pour étendre la coopération par delà les frontières notamment linguistiques.

La communauté s'élargit et se diversifie mais la question de l'hégémonie linguistique agite les pays non anglophones. Lors du premier THATCamp français en 2010, un

Manifeste pour les DH est rédigé et traduit en 11 langues. La DH2012 à Hambourg porte sur le thème « *Digital Diversity* ». Quelques mois plus tôt, une enquête en ligne, proposée en quatre langues, était lancée sur le thème « Who are you, Digital Humanists ? » afin d'interroger le spectre linguistique et géographique. L'enquête reçoit des réponses en provenance de 55 pays qui montre l'émergence de la France, de l'Espagne et de l'Amérique latine. Elle révèle surtout que l'anglais domine comme langue d'échange scientifique même si elle n'est la première langue que de moins de 23 % des répondants.

La question du respect plurilinguistique revient régulièrement dans les discussions, les protestations (par exemple, port d'un badge durant le DH2014 à Lausanne, faisant connaître les différentes langues parlées par le participant à la conférence, basculement vers le plurilinguisme avec l'espagnol dans certaines sessions, etc.) et les blogs. Des analyses ont été engagées par des participants sur le processus de sélection des contributions. Au DH2012 à Hambourg, 342 experts ont participé à l'évaluation ; leur répartition géographique montre qu'il s'agit d'un événement dominé par les collègues nord-américains et britanniques, tandis que les pays latins n'ont qu'une place marginale dans le processus de sélection des contributeurs. Pour le DH2014, l'appel à contribution, lancé en 23 langues, preuve d'ouverture à la diversité linguistique, ne reçoit d'écho qu'en anglais (92 %) ; les non-anglophones craignent peut-être une discrimination qui s'est confirmée (60 % des propositions soumises en anglais sont acceptées pour seulement 30 % dans d'autres langues).

Ces questions d'hégémonies culturelles passent aussi par les biais venant des outils (UNIX, ASCII, TEI, etc.) et

par le contenu du travail engagé, notamment la numérisation qui s'est largement faite au service de sources anglophones. Pour éviter de tomber dans le piège de ces hégémonies, des auteurs recommandent de numériser autant les traces de l'activité des personnes de couleurs et des femmes qu'on ne le fait des hommes blancs. Ils répondent ainsi aux critiques faites aux humanités numériques de manquer d'attention aux questions de race, de classe et de genre, et d'être mues par des questions de recherche plus que d'enseignement ou de problèmes de société.