

«QU'EST-CE QUE LE NUMÉRIQUE A CHANGÉ DANS NOTRE RAPPORT À LA RECHERCHE?»

Aris XANTHOS

(Université de Lausanne, Faculté des lettres,
Section des sciences du langage et de l'information)

En réponse à la question de ce que le numérique a changé à ma recherche, je pourrais presque indifféremment parler de tout *ou* de rien. J'estime à quatre-vingt-dix-huit pourcents la proportion de mes tentatives de contribuer à la Recherche, cette belle entreprise collective à laquelle nous prenons part toutes et tous, qui ont été au moins teintées et plus souvent complètement imbibées de numérique. Ma toute première publication s'appelait «ENTROPIZER 1.1: un outil informatique pour l'analyse séquentielle»¹. Si je la mentionne ici, ce n'est pas tant pour commémorer son impact – plutôt modeste – sur la communauté scientifique que pour souligner la façon dont, il y a vingt et un ans déjà, le numérique a immédiatement pris dans ma recherche le rôle central qu'il y occupe encore aujourd'hui.

Ce rôle, qui n'est assurément et heureusement pas le seul que le numérique puisse jouer dans une pratique de recherche, est de nature fondamentalement méthodologique. Qu'il s'agisse de développer des logiciels d'analyse de données textuelles ou visuelles, d'analyser quantitativement l'acquisition des langues ou d'utiliser l'intelligence artificielle pour générer interactivement des poèmes, l'informatique a toujours joué dans ma recherche le rôle d'un outil, ou plutôt d'un outillage. Dans les divers chantiers que je partage avec des collègues de la Faculté et d'ailleurs, ce que l'outillage informatique me permet de faire, c'est principalement de modéliser des phénomènes, en particulier dans le domaine des langues et littératures.

Construire un modèle ou une représentation d'un phénomène d'intérêt est une opération que beaucoup d'entre nous effectuons régulièrement dans le cadre de nos recherches. Les modèles avec lesquels je travaille se distinguent par leur caractère formel, c'est-à-dire qu'ils sont exprimés en des termes suffisamment rigoureux pour pouvoir être manipulés par des langages formels comme les mathématiques ou les langages de programmation. Parmi les avantages qu'offre le traitement formel d'un modèle, le plus évident est sans doute la possibilité d'effectuer des opérations systématiques sur des données dont le volume est souvent beaucoup plus important et la structure parfois plus complexe que ce que l'on peut envisager d'aborder sans l'informatique.

Pour n'en donner qu'un seul exemple, j'ai publié il y a dix ans un article sur le rôle de la richesse morphologique dans l'acquisition des langues, avec seize co-auteurs². Cette contribution, qui a eu plus de retentissement que ma première publication, portait sur des corpus d'acquisition dans neuf langues très diverses (du français au maya yucatèque, en passant par le finnois et le turc), chaque corpus se composant de dizaines de fichiers de transcriptions indexées chronologiquement et elles-mêmes subdivisées en tours de paroles puis en formes linguistiques codées morphologiquement. Analyser et comparer systématiquement ces données riches et complexes aurait été virtuellement impossible sans l'outil informatique.

Le recours à des modèles formels présente d'autres avantages, peut-être moins intuitivement évidents, notamment du point de vue de la communication entre les acteurs et actrices de la recherche. Mais il y a bien sûr aussi des inconvénients, dont le plus saillant est certainement que la simplification nécessairement imposée par la formalisation aux phénomènes auxquels on l'applique est parfois excessive.

On pourrait déduire de ce bref compte-rendu que le numérique n'a rien changé à ma recherche, du fait qu'il n'y avait, dans mon cas, pas de recherche avant le numérique. J'aimerais

¹ Xanthos 2000.

² Xanthos *et al.* 2011.

pourtant noter en conclusion que ma recherche a pris un vrai «tournant numérique» ces dernières années, non sur le plan des méthodes mais sur celui des objets. En contraste avec mes travaux plus anciens, les objets sur lesquels portent mes travaux récents appartiennent en effet au domaine du numérique – toujours en lien avec les sciences du langage. Je pense notamment à la recherche sur l'évolution des aspects langagiers du jeu vidéo que je mène avec Jérôme Jacquin (Section des sciences du langage et de l'information), ainsi qu'aux projets de recherche sur la communication médiée par ordinateur que je conduis dans le cadre du NCCR Evolving Language et d'un projet CROSS 2021 avec Anita Auer (Section d'anglais) et Robert West (EPFL). Comme je n'ai pas, pour autant, réformé ma méthodologie dans la direction opposée, mes méthodes et mes objets sont actuellement numériques. Et comme la communication scientifique n'a jamais été aussi souvent numérique qu'aujourd'hui, il n'y a sans doute pas lieu de s'inquiéter d'un manque de cohérence entre forme et fond!

RÉFÉRENCES BIBLIOGRAPHIQUES

- XANTHOS A., 2000: «ENTROPIZER 1.1: un outil informatique pour l'analyse séquentielle», in *Actes des 5es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2000)*. Lausanne: EPFL: 357-364.
- XANTHOS A., LAAHA S., GILLIS S., STEPHANY U., AKSU-KOÇ A., CHRISTOFIDOU A., GAGARINA N., HRZICA G., KETREZ F.N., KILANI-SCHOCH M., KORECKY-KRÖLL K., KOVAČEVIĆ M., LAALO K., PALMOVIĆ M., PFEILER B., VOEIKOVA M.D., DRESSLER W.U., 2011: «On the role of morphological richness in the early development of noun and verb inflection», in *First Language*, 31(4): 461-479.