

VI. Conclusion

Ainsi, le travail de recensement des enquêtes quantitatives sur les langues parlées en France a abouti sur une liste de plus de 100 enquêtes, présentant une grande hétérogénéité, à la fois en terme de territoires représentés, de types de commanditaires, mais aussi d'échantillonnage et de qualité des données produites. On observe tout d'abord que certaines langues régionales et certains territoires sont surreprésentés (ouest et sud-ouest de la France, en particulier), tandis que le nombre d'enquêtes traitant des langues migratoires est extrêmement limité. Concernant les Outre-mer, il y a également une grande hétérogénéité des situations, avec d'un côté les COM du Pacifique produisant de nombreuses données linguistiques à travers leurs instituts statistiques locaux, et de l'autre les DOM sur lesquels un certain nombre d'enquêtes sont effectués par la statistique publique mais avec de nombreuses difficultés méthodologiques, en particulier sur les dénominations, qui affectent parfois l'exploitation des données. On note également que les seuls territoires de France où il existe une absence totale de données linguistiques sont les COM de Saint Barthélemy, Saint Pierre-et-Miquelon, et Saint-Martin.

Malgré un grand nombre d'enquêtes recensées, il est néanmoins très difficile d'accéder aux données issues de celles-ci, car il existe un phénomène fort de perte des données, causé à la fois par l'ancienneté de ces enquêtes, l'absence ou le manque de connaissance des structures de stockage et d'archivage, ainsi que le recours fréquent à des prestataires privés qui ne livrent pas systématiquement les bases de données à l'issue des enquêtes (ou bien les commanditaires ne les demandent pas).

En conséquence, pour les acteurs publics comme la société civile, il existe un réel problème d'accès aux données sur les langues en France, qui amène dans certains cas à utiliser des sources douteuses comme les sites sur les langues réalisés par les organisations missionnaires américaines, à défaut de pouvoir trouver d'autres informations ailleurs. Dans le cadre de la loi Numérique de 2016, ainsi que des enjeux de la science ouverte⁶⁹, il est impératif de travailler vers une ouverture des données pour permettre leur utilisation et leur exploitation à la fois par le monde de la recherche, les acteurs publics et la société civile.

⁶⁹ <https://www.ouvrirlascience.fr/accueil/>