

Délégation générale à la langue française et aux langues de France

ENCOURAGER LE DÉVELOPPEMENT DES TECHNOLOGIES DE LA LANGUE

I. Accompagner l'effort de recherche en faveur des technologies de la langue

La Délégation générale à la langue française et aux langues de France s'intéresse de près à l'innovation et à la recherche dans le domaine de la linguistique, **domaine d'excellence pour la France** qui compte un très grand nombre d'unités de recherche sur ce domaine. À ce titre, **les technologies de la langue** se révèlent être d'une importance capitale. On entend par technologies de la langue toutes les techniques permettant de manipuler, transformer, synthétiser la langue, qu'elle soit écrite, orale, ou audiovisuelle.

Les technologies les plus connues permettent notamment la traduction dite « automatique » et la synthèse vocale, mais il y a bien d'autres applications : aide à la traduction, sous-titrage automatique, aide à la rédaction (correcteurs orthographiques et grammaticaux), reconnaissance vocale, indexation automatique de documents...

Elles constituent **un enjeu culturel et linguistique considérable** : en effet, plus une langue est facilement accessible, traduisible et manipulable au moyen de ces technologies, plus elle sera couramment utilisée sur les réseaux numériques. Il importe donc de s'assurer de la bonne place du français et des langues de France dans le développement de ces technologies, en soutenant la recherche, en favorisant les partenariats public-privé, en participant aux travaux et aux études financées par la Commission européenne.

Du point de vue des technologies, **les problèmes d'ingénierie linguistique sont loin d'être résolus** et demanderont encore pour très longtemps des efforts importants de recherche : la disponibilité d'applications grand public qui fonctionnent de manière très imparfaite ne signifie pas que les problèmes sont réglés mais plutôt qu'en dépit de leurs imperfections les applications d'aujourd'hui (traduction automatique, SIRI, analyse de sentiments ou d'opinions, etc) fournissent déjà des services appréciables et génèrent une forte activité économique.

Raison de plus pour continuer les recherches dans le sens d'une part d'un **enrichissement des analyses et des représentations sémantiques** ; d'autre part d'une **ouverture vers des contenus plus difficiles à traiter** que des textes bien écrits (écrits dans les médias sociaux, transcription d'audio et de vidéo, etc.)

Du point de vue politique / sociétal, les technologies de la langue (par exemple la traduction, la correction automatique, les outils d'aide à la recherche d'information ou les aides à la compréhension) sont des outils essentiels pour **favoriser « l'inclusion numérique »** de larges parties de la population dans un contexte où bien lire et écrire au moins une langue est indispensable pour tirer le meilleur parti des sources d'information et des connaissances disponibles en ligne. Il reste ici beaucoup à faire, y compris autour des outils d'aides à l'apprentissage des langues.

II. Proposer un programme en faveur des technologies de la langue

Un gros travail reste également à accomplir pour **documenter et outiller les langues de France**¹ : il s'agit bien sûr des langues régionales, telles que par exemple l'occitan, le breton, le corse ou les langues créoles, mais aussi d'autres langues non territoriales telle que la langue des signes française. Bien que certaines initiatives existent dans le domaine des technologies, notamment en ce qui concerne le breton ou l'occitan, elles disposent de très peu de moyens de recherche.

Or, ces technologies (synthèse et reconnaissance vocales, correction orthographique, traduction automatique...) sont nécessaires pour assurer, dans certains cas, **la simple survie de ces langues**, dans d'autres cas leur présence dans la vie sociale. **Un travail d'identification** de ces ressources est en cours, mais il doit être prolongé par **une action publique vigoureuse**, en liaison avec les collectivités locales et l'ensemble des acteurs académiques et associatifs concernés, pour équiper les langues de France en technologies de la langue.

Une action analogue au programme « **Technolangue** » (qui, doté d'un budget d'environ un million d'euros, associait en 2003-2005 les ministères de la Culture, de l'Industrie, et de la Recherche) pourrait être envisagée pour développer les ressources structurées et les technologies de la langue pour le français et les langues de France.

Ce programme, dont l'utilité sur la croissance, l'emploi et, plus généralement, **sur l'économie française** est soulignée dans le rapport intitulé « La francophonie et la francophilie, moteurs de croissance durable » remis par Jacques Attali à la présidence de la République en août 2014². Il pourrait être cofinancé dans le cadre du **Programme d'Investissements d'Avenir (PIA)**.

¹Une liste des langues officiellement reconnues comme « Langues de France » par la délégation figure à cette adresse : http://www.culture.gouv.fr/culture/dglf/lgfrance/lgfrance_presentation.htm

²Rapport « La francophonie et la francophilie, moteurs de croissance durable », Jacques ATTALI, 2014 :

http://www.ladocumentationfrancaise.fr/rapports-publics/144000511/index.shtml#U_xa_fcxRqk.twitter?xtor=AL-1877