
L'esperanto, langue pivot : DLT de BSO

A Utrecht (Pays-Bas) depuis 1983, le Buro voor Systeemontwikkeling (BSO) développe le projet de traduction automatique DLT (Distributed Language Translation) qui sera opérationnel vers 1992 et qui utilise l'espéranto comme langue pivot.

Il s'agit de permettre aux utilisateurs de recevoir dans la langue de leur choix des documents techniques distribués par télématique. L'idée de départ est qu'un logiciel de traduction vraiment performant et entièrement automatique n'est pas réalisable d'ici 1992, si bien que l'utilisation d'une langue pivot se justifie à double titre : non seulement elle diminue considérablement le nombre de logiciels nécessaires lorsque plusieurs langues sources et plusieurs langues cibles sont concernées, mais en outre, elle peut être conçue de telle sorte que sa phase de traduction vers la langue cible soit, elle, entièrement automatique, alors que la traduction de la langue source vers la langue pivot nécessite l'intervention d'un homme et non d'un traducteur, car seule la connaissance de la langue source est requise.

L'utilisateur francophone, à l'aide d'un logiciel de traduction du français vers la langue pivot et de la langue pivot vers le français, pourra donc distribuer ses documents dans n'importe quelle autre langue et recevoir en français des documents saisis dans n'importe quelle autre langue : le serveur télématique ne gèrera que des textes en langue pivot, la traduction vers celle-ci se fera en temps réel au moment de la saisie grâce à quelques questions, posées automatiquement par le système à l'utilisateur, dans sa propre langue, ou permettant de lever toutes les ambiguïtés du texte saisi, et la traduction de la langue pivot vers la langue cible se fera elle aussi chez l'utilisateur, à la réception du texte. En principe, un texte saisi en français et appelé en français, s'il transite par le système DLT, subira lui aussi une double traduction : français-espéranto et espéranto-français.

Le choix de l'espéranto comme langue pivot résulte de plusieurs considérations : cette langue, construite il y a un siècle, est grammaticalement plus

régulière que toutes les autres langues nationales, ce qui facilite son traitement automatique. Elle existe et fonctionne déjà, ce qui évite la perte de temps que représenterait la création d'un code artificiel remplissant les conditions requises pour une langue pivot. Un tel code artificiel ne serait pas autonome, dans la mesure où les divers collaborateurs du projet continueraient à penser dans leur propre langue lorsqu'ils utiliseraient ce code (il y aurait donc autant de versions distinctes de ce code que de langues en présence), alors que les collaborateurs de DLT pensent en espéranto.

Comme la langue pivot ne sera pas accessible à l'utilisateur lorsque le système sera opérationnel, il n'est pas nécessaire qu'elle ait la même souplesse qu'une langue vivante et une version quelque peu modifiée de l'espéranto a été élaborée à cette fin, n'utilisant qu'une partie des structures syntaxiques possibles et créant certains suffixes et certains termes spéciaux pour éviter les ambiguïtés, mais l'expérience prouve que les modifications nécessaires sont moins importantes que celles envisagées initialement. Il ne s'agit pas, en effet, d'élaborer un "système linguistique universel", qui éviterait la phase de transfert entre la langue source et la langue cible, mais seulement une langue suffisamment claire pour que la traduction de la langue pivot vers la langue cible puisse se faire sans intervention humaine, et cette condition est déjà presque vérifiée par l'espéranto tel qu'il est effectivement pratiqué aujourd'hui.

A l'heure actuelle, des informaticiens et des linguistes, spécialistes de l'anglais, du français et de l'espéranto, de diverses nationalités, sont réunis à Utrecht pour la première étape de ce projet DLT (traduction entre l'anglais et le français), mais le travail fait sur l'anglais et le français ne tardera pas à être généralisé à d'autres langues, en collaboration avec des équipes localisées dans d'autres villes d'Europe ou du reste du monde ... L'analyse syntaxique et l'analyse sémantique sont développées parallèlement, un système d'apprentissage permettant de réduire le nombre d'ambiguïtés que le système ne peut résoudre sans intervention humaine est également envisagé et plusieurs ouvrages ont été publiés sur ce projet qui a été présenté dans un certain nombre de colloques. ■