

Communication

Les mots
pour le dire



Langage



L'équipe de l'UNamur qui a développé le dictionnaire, porte d'entrée pour accéder au monde des sourds et malentendants.

© UNAMUR

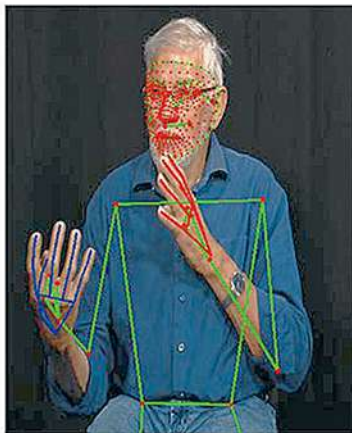
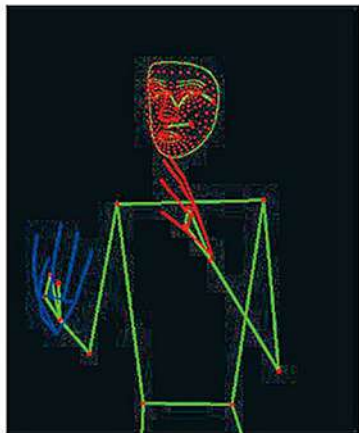
C'est une première mondiale. Un dictionnaire langue des signes-langue vocale vient de sortir. Il est le fruit de travaux de recherche en linguistique et en intelligence artificielle à l'UNamur.

Par Laetitia Theunis

En Belgique, la langue des signes est pratiquée par quatre mille personnes. Jusqu'ici, aucun dictionnaire ne lui était consacrée, vers aucune langue vocale. C'est chose faite : des chercheurs en linguistique et en informatique de l'UNamur viennent de mettre en ligne un *Dictionnaire langue des signes de Belgique francophone – français*. Soutenu financièrement par le Fonds Baillet Latour, il a été conçu à l'intention des familles entendant avec un enfant sourd, des locuteurs d'une langue signée qui doivent passer au français, plus largement de toute personne ayant besoin occasionnellement de faire le pont entre français et langue des signes, ainsi qu'à l'intention des classes bilingues, lesquelles ont mis en œuvre ce projet.

« Jusqu'à la naissance de ce premier dictionnaire langue des signes-français, c'étaient nous, enseignants, qui devions jouer le rôle de répertoire pour les enfants. La langue des signes ne s'écrivant pas, nous devions réaliser des vidéos présentant le signe d'un nouveau mot, le mettant en contexte, et joindre ces vidéos par QR code aux feuilles distribuées en classe. C'était un travail monumental », explique Magaly Ghesquière, coordinatrice des classes bilingues langue des signes-français et enseignante à la Communauté scolaire Sainte-Marie à Namur. Là, de petits groupes d'enfants sourds sont insérés dans des classes d'enfants entendants de l'enseignement général. « Pour les premiers, le dictionnaire permet, par exemple, la compréhension quasi immédiate ...

Pour que l'IA puisse reconnaître les signes, une sorte de squelette numérique est collé sur chaque signeur.



© UNAMUR

... d'expressions complexes dans des romans écrits en français. Pour les seconds, en les aidant à acquérir les signes nécessaires pour converser avec leurs camarades sourds, il est la clé qui ouvre la porte d'un autre monde. » Celui du silence.

COMMENT ÇA FONCTIONNE

Le matériel nécessaire se compose d'un ordinateur ou d'une tablette ou d'un smartphone, équipé d'une caméra et connecté à Internet avec un bon débit. Ensuite, il faut rejoindre le site Web du dictionnaire.

Pour qui veut connaître le signe correspondant à un mot français (1), il suffit d'introduire ce dernier pour qu'une vidéo s'anime, montrant une personne effectuant le signe. Des exemples, permettant à l'utilisateur de voir le signe utilisé dans son contexte, sont une précieuse aide. « En effet, en langue des signes comme dans toutes les langues du monde, il n'y a pas de transposition directe d'un mot en un autre. Le contexte joue une part importante », précise Laurence Meurant, directrice du Laboratoire de langue des signes de Belgique francophone à l'UNamur.

Le dictionnaire est également opérationnel dans l'autre sens linguistique (2). Face à la webcam, en positionnant son visage dans le cercle en pointillés, l'utilisateur saisit un signe et la plateforme affiche sa traduction en français. « Pour 80 % des requêtes, le signe correct est dans le top 5 proposé par l'intelligence artificielle (IA) », se réjouit Jérôme Fink, doctorant en informatique, cheville ouvrière de ce projet depuis cinq ans.

SQUELETTE NUMÉRIQUE

Pour parvenir à une telle prouesse, les chercheurs en intelligence artificielle se sont basés sur un riche corpus de plus de 36 heures de vidéos de signes, annotées et analysées une par une par des signeurs. Il comprend quelque 180 000 signes, représentant

4 600 signes répétés de multiples fois par une centaine de signeurs. Un travail de bénédictin qui a occupé les linguistes de l'UNamur pendant près de vingt ans. « Une base de données suffisamment large était primordiale pour que l'IA puisse parvenir à reconnaître un signe devant n'importe quelle webcam », précise Laurence Meurant.

« Pour aider l'IA dans cette tâche, nous avons collé sur chaque signeur, de façon informatique et automatique, une sorte de squelette numérique composé de points placés sur les éléments corporels utilisés en langage des signes : menton et bouche du signeur, contours de son visage, commissure de ses yeux, mais aussi sur ses coudes, ses mains, ses doigts », détaille Benoît Frenay, professeur au sein du Namur Digital Institute. Grâce à ce système, l'IA connaît, à chaque moment, la position de chaque point corporel impliqué dans le signe à reconnaître. Pour affiner sa compétence, elle s'est entraînée, et continue à le faire, sur les dizaines de milliers de vidéos de signes réalisées par des signeurs à la dextérité et à la morphologie particulières, avec des couleurs et des lumières différentes.

« Pour que l'IA puisse identifier un signe et transmettre sa traduction française, il lui fallait s'entraîner sur un minimum de trente occurrences de chaque signe, ajoute Anthony Cleve, copilote de ce projet avec Laurence Meurant. Dans le corpus, c'est le cas pour pas moins des 750 signes les plus couramment utilisés par les signeurs. Soit 35 % de leur lexique de base. Dans le sens français vers langue des signes, la tâche était bien plus aisée et quelque 2 100 signes de noms communs et de verbes sont d'ores et déjà accessibles. »

Les prochaines années verront cette banque de données s'enrichir et l'IA se bonifier. De quoi favoriser l'inclusion des personnes sourdes et malentendantes dans la société. ▀

(1) corpus-lsfb.be

(2) dico.corpus-lsfb.be

Un travail de bénédictin qui a occupé les linguistes de l'UNamur pendant près de vingt ans.