



# UN DICTIONNAIRE QUI DONNE LA PAROLE AU MONDE DU SILENCE

Publié le 26 octobre 2022



par Laetitia Theunis

C'est une première mondiale. [Le tout premier dictionnaire langue des signes – langue vocale vient de sortir](#). Il est le fruit d'une vingtaine d'années de recherche en langue des signes francophone de Belgique et d'une riche collaboration de 4 ans entre chercheurs en linguistique et en informatique à l'UNamur. Soutenu financièrement par le [Fonds Baillet Latour](#), il a été conçu à destination des classes bilingues, des familles entendant avec un enfant sourd, des locuteurs d'une langue signée qui doivent passer au français, mais aussi de toute personne qui a besoin occasionnellement de faire le pont entre français et langue des signes. De quoi favoriser l'inclusion des personnes sourdes et malentendantes dans la société.

## Dans les deux sens

Concrètement, comment fonctionne ce dictionnaire ? Comme matériel, rien d'extravagant n'est nécessaire : il suffit d'avoir un ordinateur ou une tablette ou un smartphone, équipé d'une caméra et connecté à Internet avec un bon débit.

Ensuite, il suffit d'[introduire un mot en français pour qu'une vidéo révèle le signe correspondant](#). Des exemples contextuels, 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.

Le dictionnaire fonctionne également dans l'autre sens linguistique. Face à la webcam, qu'importe l'endroit où il se trouve, l'utilisateur saisit un signe en langue des signes et la plateforme affiche la traduction en français. « Pour 80 % des requêtes, le signe correct est dans le top 5 proposé par l'intelligence artificielle », se réjouit Jérôme Fink, doctorant en informatique, cheville ouvrière de ce projet depuis 5 ans. Une liste de suggestions d'usage du signe, sous forme de vidéos enrichit et affine la recherche.

## Un squelette numérique

« [Passer du français à la langue des signes](#), c'était déjà un fameux défi technique. [Passer de la langue des signes au français](#) fut bien plus complexe encore », explique Benoît Frenay, professeur au sein de l'institut NADI ([Namur Digital Institute](#)), et spécialiste en intelligence artificielle.

En effet, d'une part, chaque signeur est unique. Certains enchaînent les mouvements plus rapidement que d'autres, sont plus aguerris, ont une morphologie particulière. D'autre part, chaque vidéo est unique, avec sa propre luminosité, avec des couleurs de vêtements et de peaux particulières. Chacun de ces paramètres complexifie la tâche de l'intelligence artificielle.

« Pour l'aider, 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. Ils sont tous liés à la partie haute du corps. Ainsi, des points numériques sont placés sur le menton du signeur, sur les contours de son visage, à la commissure de ses yeux, sur sa bouche, mais aussi sur ses coudes, ses mains, ses doigts », poursuit-il.



Grâce à ce système, l'IA connaît, à chaque moment, la position de chaque point corporel impliqué dans le signe à reconnaître.



Squelette numérique sur un signeur © UNamur

## **20 années de travail de fourmi**

Sans le travail monumental réalisé en amont par la Pre Laurence Meurant, cette prouesse informatique, et par là le dictionnaire, n'aurait pu voir le jour. Depuis 20 ans, son équipe du LSFb-Lab ([Laboratoire de langue des signes de Belgique francophone](#)), au sein de l'[institut NaLTT](#), travaille sur la langue des signes de Belgique francophone et a créé un riche corpus de 36h37 de vidéos annotées et analysées une par une par des signeurs. Un travail de bénédictin. D'autres chercheurs ont traduit ces vidéos en français.

C'est sur cette base de données bilingue que les chercheurs en informatique ont pu construire les algorithmes d'intelligence artificielle utilisés dans la plateforme du dictionnaire. « Une base de données suffisamment large était primordiale pour que l'IA puisse parvenir à reconnaître un signe devant n'importe quelle webcam », explique-t-elle.

Au total, le corpus vidéo comprend quelque 180.000 signes, représentant 4600 signes répétés de multiples fois par une centaine de signeurs.

## **35 % du lexique courant des signeurs**

Parmi ces 4000 signes, la moitié est composée de noms propres et de noms peu courants. Les chercheurs se sont focalisés sur l'autre moitié du corpus, reprenant les noms communs et les verbes.

« Parmi ceux-ci, pour que l'intelligence artificielle puisse identifier un signe et transmettre sa traduction française, il lui fallait s'entraîner sur un minimum de 30 occurrences de chaque signe. C'est le cas pour pas moins des 750 signes les plus couramment utilisés par les signeurs. Soit 35 % de leur lexique courant. Dans le sens français vers langue des signes, la tâche était bien plus aisée et les 2100 signes sont d'ores et déjà accessibles », précise Anthony Cleve, professeur en informatique à l'institut NADI, pilote de ce projet aux côtés de Laurence Meurant.

Et celle-ci d'ajouter, « à l'avenir, ce dictionnaire va aussi permettre de grandes avancées du côté linguistique. Car, à partir de maintenant, l'outil de reconnaissance des signes va nous permettre d'annoter signe par signe beaucoup plus rapidement que précédemment. Autrement dit, ce qui se faisait avant manuellement, va désormais pouvoir être soutenu par la reconnaissance automatique de l'ordinateur. »

« La bibliothèque à partir de laquelle on travaille va pouvoir s'étendre beaucoup plus rapidement. De la sorte, les recherches linguistiques vont pouvoir être plus solides, plus variées. Cet outil va aussi permettre de développer, pour les traducteurs et interprètes, des outils d'aide à la traduction. »

## **Implication des utilisateurs**

La langue des signes francophone de Belgique est pratiquée par environ 4.000 personnes. L'idée de la création d'un dictionnaire bilingue prend sa source dans la collaboration qu'entretient depuis près de 20 ans, l'UNamur avec les classes bilingues (français – langue des signes) fondées par Ecole et Surdit   à Sainte-Marie Namur.

C'est pourquoi « il   tait crucial que le dictionnaire, m  me s'il s'agit d'une innovation, soit utile, facile    comprendre et    prendre en main par les utilisateurs finaux », indique Pr Bruno Dumas, vice-pr  sident de l'institut NADI.



Pour garantir l'intuitivité du site, des signeurs, des experts et des enseignants ont été impliqués à différents stades du développement du projet.

Ainsi, dans le cas où quelqu'un fait un signe devant sa caméra afin qu'il soit reconnu par l'IA, celle-ci avait besoin que l'utilisateur démarre sa séquence vidéo avec une pose bien précise. Il a été décidé de rendre cela ludique, sous la forme d'un photomaton : avec un cercle pour positionner son visage et un minuteur.

Souhaitons bon vent à cet outil gratuit qui vise davantage d'inclusion des personnes sourdes et malentendantes dans la société.