



Un modèle neuro- computationnel décode la parole en la prédisant

Des scientifiques de l'UNIGE ont mis au point un modèle neuro-informatique, alliant oscillations neuronales et codage prédictif, qui permet d'expliquer comment le cerveau identifie les syllabes dans la parole naturelle.

L'analyse du langage par le cerveau passe par la reconnaissance des syllabes. Des scientifiques de l'Université de Genève (UNIGE) et du Pôle de recherche national «Evolving Language» ont conçu un modèle neuro-computationnel destiné à reproduire le mécanisme complexe qui est mis en œuvre par le système nerveux central pour mener à bien cette opération. Réunissant deux cadres théoriques indépendants, le modèle utilise, d'une part, l'équivalent des oscillations neuronales produites par l'activité cérébrale pour décortiquer le flux sonore. De l'autre, il fonctionne en accord avec la théorie dite du codage prédictif selon laquelle le cerveau optimise sa perception en prédisant en permanence le sens des signaux qu'il perçoit. Le programme qui en résulte, présenté dans la revue *Nature communications*, a permis de reconnaître en direct des milliers de syllabes contenues dans des centaines de phrases parlées en langage naturel, validant ainsi l'idée que les oscillations neuronales servent à coordonner le flux de syllabes entendues avec les prédictions faites par notre cerveau.

«L'activité cérébrale produit des oscillations neuronales que l'on peut mesurer notamment à l'aide de l'électroencéphalographie, explique Anne-Lise Giraud, professeure au Département des neurosciences fondamentales de la Faculté de médecine de l'UNIGE et codirectrice du PRN «Evolving Language». Il s'agit d'ondes électromagnétiques résultant de l'activité électrique et cohérente de réseaux entiers de neurones. Il en existe plusieurs sortes, définies selon leur fréquence. On les appelle les ondes alpha, bêta, delta thêta ou encore gamma. Pris individuellement ou superposés, ces rythmes sont associés à différentes fonctions cognitives, tel que la perception, la mémoire, l'attention, la vigilance etc.»

Bien que les oscillations neuronales soient associées à ces fonctions, les neuroscientifiques ignorent encore si elles y contribuent activement et de quelle façon. Dans une étude antérieure parue en 2015, l'équipe d'Anne-Lise Giraud a déjà montré que les ondes thêta (de basse fréquence) et gamma (haute fréquence) se coordonnent pour séquencer le flux sonore en syllabes et analyser leur contenu de manière à les reconnaître.

Le rythme des syllabes

Les scientifiques genevois-es ont mis au point un premier modèle neuro-computationnel inspiré de ces rythmes physiologiques, dont les performances de séquençage des syllabes en direct (on-line) se sont révélées supérieures à celles des systèmes traditionnels de reconnaissance automatique de la parole.

Dans ce premier modèle, les ondes thêta (comprises en 4 et 8 Hertz) permettent de suivre le rythme des syllabes au fur et à mesure qu'elles



© Carla Da Silva

Anne-Lise Giraud

Illustrations haute définition

sont perçues par le système. Les ondes gamma (autour de 30 Hertz), servent à découper le signal auditif en tranches plus petites et à l'encoder. Cela produit un profil «syllabique» associé à chaque séquence sonore qui peut être comparé, a posteriori, à une bibliothèque de syllabes connues. L'un des avantages d'un tel modèle, c'est qu'il peut s'adapter spontanément à la vitesse de parole qui peut varier d'un individu à l'autre.

Codage prédictif

Dans ce nouvel article, et pour coller encore plus à la réalité biologique, Anne-Lise Giraud et son équipe ont développé un nouveau modèle dans lequel ils ont intégré des éléments d'un autre cadre théorique, indépendant des oscillations neuronales, qui est celui du «codage prédictif».

«Selon cette théorie, si le cerveau fonctionne de façon si optimale, c'est qu'il tente en permanence d'anticiper ce qui se passe dans l'environnement, explique Itsaso Olasagasti, chercheur dans l'équipe d'Anne-lise Giraud, qui a supervisé la mise en œuvre du nouveau modèle. Dans le cas du langage, il tenterait, à mesure que la parole se déroule, de s'expliquer les causes les plus probables des sons perçus par l'oreille sur la base d'un jeu de représentations mentales apprises, et mises à jour en permanence.»

«Nous avons donc développé un modèle informatique simulant ce codage prédictif, explique Sevada Hovsepyan, chercheur au Département des neurosciences fondamentales et premier auteur de l'article. Et nous l'avons implémenté en y incorporant des mécanismes oscillatoires.»

Testé sur 2888 syllabes et 220 phrases

Le son entrant dans le système est d'abord modulé par une onde thêta (lente) qui ressemble à ce que produisent les populations de neurones. Elle permet de signaler les contours des syllabes. Ensuite, des trains d'ondes gamma (rapides) permettent d'encoder la syllabe au fur et à mesure qu'elle est perçue. Au cours du processus, le système propose des syllabes possibles et corrige son choix si nécessaire. Après quelques aller-retours entre les deux niveaux, il découvre la bonne syllabe. Le système est ensuite mis à zéro à la fin de chaque syllabe perçue.

Le modèle a été testé avec succès à l'aide de 2888 syllabes différentes, contenues dans 220 phrases, prononcées en langage naturel et en anglais. «Nous avons réussi d'une part à réunir dans un seul modèle informatique deux cadres théoriques très différents, explique Anne-Lise Giraud. De l'autre, nous avons montré que les oscillations neuronales permettent très probablement d'aligner rythmiquement le fonctionnement endogène du cerveau avec les signaux venus de l'extérieur via les organes sensoriels. Si on replace cela dans la théorie du codage prédictif, cela signifie que ces oscillations permettent probablement au cerveau d'émettre la bonne hypothèse exactement au bon moment.»

contact

Anne-Lise Giraud

Professeure ordinaire
Département des neurosciences
fondamentales
Faculté de médecine
Codirectrice du Pôle de re-
cherche national «Evolving
Language»
anne-lise.giraud@unige.ch
+41 22 379 55 47

DOI: 10.1038/s41467-020-16956-5

UNIVERSITÉ DE GENÈVE Service de communication

24 rue du Général-Dufour
CH-1211 Genève 4
Tél. +41 22 379 77 17
media@unige.ch
www.unige.ch