

Mémoire de fin d'études
Pour l'obtention du diplôme d'Ingénieur d'État en Informatique
Option : Systèmes Informatiques

Création d'un corpus de l'aphasie de Broca et
développement d'un système Speech-to-speech de
réhabilitation de la parole

Réalisé par :
BELGOUMRI Mohammed
Djameleddine
im_belgoumri@esi.dz

Encadré par :
Pr. SMAILI Kamel
smaili@loria.fr
Dr. LANGLOIS David
david.langlois@loria.fr
Dr. ZAKARIA Chahnez
c_zakaria@esi.dz

Table des matières

Page de garde	1
Table des matières	2
Table des figures	3
Sigles et abréviations	4
1 Aphasie de Broca	5
1.1 Notes d'histoire	5
1.2 Généralités sur le cerveau	6
1.3 Types d'aphasie	7
1.4 Étiologie et épidémiologie de l'aphasie	8
2 Apprentissage séquence à séquence	9
2.1 Énoncé du problème	9
2.2 Réseaux de neurones feed-forward	10
2.2.1 Généralités	10
2.3 Réseaux de neurones récurrents	11
3 Traduction automatique	12
3.1 Classification des méthodes de traduction automatique	12

3.2	Traduction automatique à base de règles	13
3.3	Traduction automatique statistique	13
3.4	Traduction automatique neuronale classique	14
3.5	Traduction automatique à base de transformeurs	14
Bibliographie		17

Table des figures

1.1	Cerveau de Victor Louis Leborgne avec la lésion encadrée	6
1.2	Encéphale humain	7
1.3	Division morphologique et fonctionnelle du cerveau.	7
1.4	Classification de certains types d'aphasie	8
2.1	Architecture sous-jacente d'un FFN (réseau de neurones feed-forward) de profondeur 3.	10
3.1	Taxonomie des méthodes de traduction automatique.	13
3.2	Triangle de Vauquois	13
3.3	L'architecture du transformeur	15

Sigles et abréviations

RNN (réseau de neurones récurrent) de l'anglais : recurrent neural network

Chapitre 1

Aphasie de Broca

L’aphasie ; emprunté au Grec ancien ἀφασία qui veut dire “mutisme”, est un trouble de communication d’origine neurologique « Dictionnaire de français Larousse », s. d. Elle affecte la capacité à comprendre le langage, s’y exprimer ou les deux. L’aphasie n’est pas causée par un trouble moteur, sensoriel, psychique ou intellectuel CHAPEY, 2008. Sa cause principale est un AVC, mais elle peut également être le résultat d’une infection ou tumeur cérébrale, un traumatisme crânien, un trouble métabolique comme le diabète ou une maladie neurodégénérative comme l’Alzheimer HALLOWELL, 2017.

1.1 Notes d’histoire

Louis Victor Leborgne, né en 1809 à Moret-sur-Loing commença à perdre la capacité de parler à l’âge de 30 ans. Il fut admis à l’hôpital de Bicêtre où il passerait 21 ans pendant lesquelles, il ne communiquait qu’en produisant le son “tan”, typiquement répété deux fois, si bien qu’on lui a donné le surnom “monsieur Tantan” MOHAMMED et al., 2018.

Le 11 avril 1861, monsieur Leborgne fut examiné par Dr. Pierre Paul Broca pour une gangrène dans son pied droit. Dr. Broca s’intéressa au trouble linguistique dont souffrait son patient LORCH, 2011. Il fit l’observation que les facultés intellectuelles et motrices de monsieur Leborgne étaient intactes, il en conclut qu’elles ne peuvent être à l’origine de son handicap. Broca nomma “aphémie” ce type de situation BROCA, s. d., il en écrivit :

“Cette abolition de la parole, chez des individus qui ne sont ni paralysés ni idiots, constitue un symptôme assez singulier pour qu’il me paraisse utile de la désigner sous un nom spécial. Je lui donnerai donc le nom d’aphémie (α privatif ; φημι, je parle, je prononce) ; car ce qui manque à ces malades, c’est seulement la faculté d’articuler les mots.”

— BROCA, s. d.

Dr. Broca prit ce constat comme confirmation de ce qu’il appelait “le principe de localisations cérébrales”. Il s’agit de l’idée que ce dernier fonctionne comme système à plusieurs composants plutôt qu’un monolithe et que les fonctions cognitives sont spatialement localisées FODOR, 1983.

Quand monsieur Leborgne mourut le 17 avril, Dr. Broca lui fit l'autopsie. En ouvrant le crâne, il observa une lésion dans le cortex inférieur gauche du lobe frontale (Voire Figure 1.1). Il en déduit (1) que cette lésion était à l'origine de l'aphémie de monsieur Leborgne et (2) que la partie affectée du cerveau est responsable d'articuler des expressions dans le langage BROCA, s. d. ; LORCH, 2011 ; MOHAMMED et al., 2018.

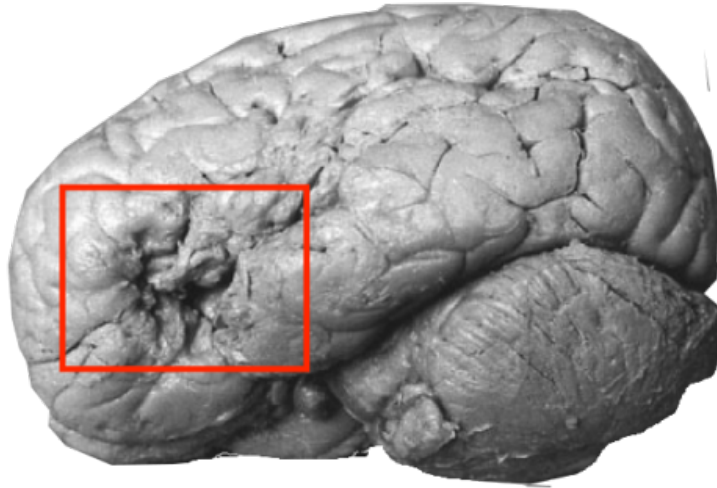


FIGURE 1.1 – Cerveau de Victor Louis Leborgne avec la lésion encadrée

1.2 Généralités sur le cerveau

Pour mieux comprendre l'aphasie en général et celle de Broca en particulier, il convient de commencer avec le cerveau. Il s'agit du système biologique le plus complexe connu. Avec le cervelet et le tronc cérébral, il forme l'encéphale (voir Figure 1.2). Le cerveau se charge du traitement des flux nerveux sensoriels et moteurs. Il est aussi le siège des hautes fonctions cognitives comme l'inférence logique, l'émotion et — crucialement pour notre étude — le traitement du langage.

Le cerveau est composé de deux hémisphères ; chacun desquels se divise en lobes : frontal, temporal, pariétal et occipital (voir Figure 1.3a). La surface du cerveau s'appelle le "cortex cérébral". Il présente plusieurs circonvolutions qui augmentent considérablement sa surface. Le cortex cérébral est divisé en régions fonctionnelles que nous appelons "aires" (voir Figure 1.3b). Le travail de Dr. Broca sur le cas de M. Leborgne est largement reconnu comme l'origine de cette division.

Une autre division importante et due à l'anatomiste Allemand Korbinian Brodmann. Elle se base sur l'organisation cellulaire des neurones pour segmenter le cortex cérébral en 52 régions aussi nommées "aires". En dépit d'avoir une définition morphologique, les aires de Brodmann sont largement alignés sur les aires fonctionnelles de la Figure 1.3b BRODMANN, 2007.

Dans cette étude, nous donnons un intérêt particulier aux aires de Brodmann 44, 45 et 39, 40. En effet, les deux premiers correspondent à la région du cerveau de M. Leborgne

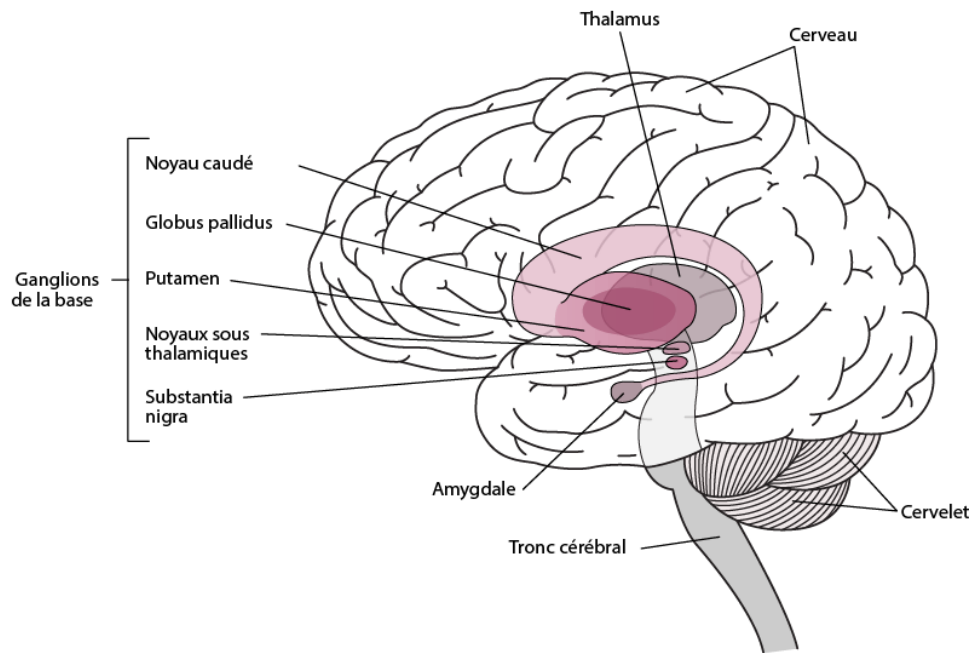


FIGURE 1.2 – Encéphale humain

où Dr. Broca trouva la lésion. Elles portent donc son nom : aire de Broca.

1.3 Types d'aphasie

La définition qu'on a donnée de l'aphasie s'applique à une multitude de troubles qui touchent différents aspects de la communication HALLOWELL, 2017, p. 135, 136. De ce fait, une classification des aphasies a été établie sur la base de leurs effets.

Plus spécifiquement, on classe une aphasie selon si elle touche l'une des trois tâches suivantes : parler couramment, comprendre le langage et répéter la parole. Cela donne

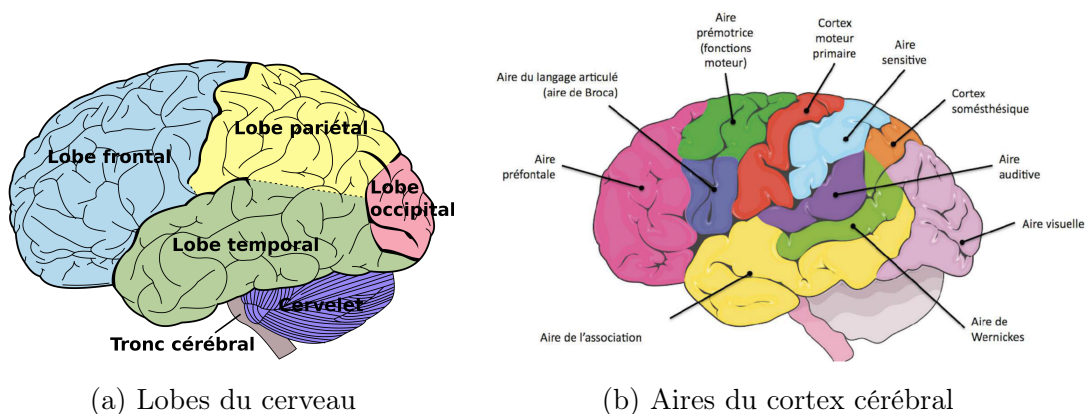


FIGURE 1.3 – Division morphologique et fonctionnelle du cerveau.

lieu aux huit classes qu'on voit sur la Figure 1.4. On note bien que cette classification n'est pas complète, l'aphasie primaire progressive par exemple n'y est pas. Cependant, elle reste utile pour étudier les types d'aphasie qui y sont présents.

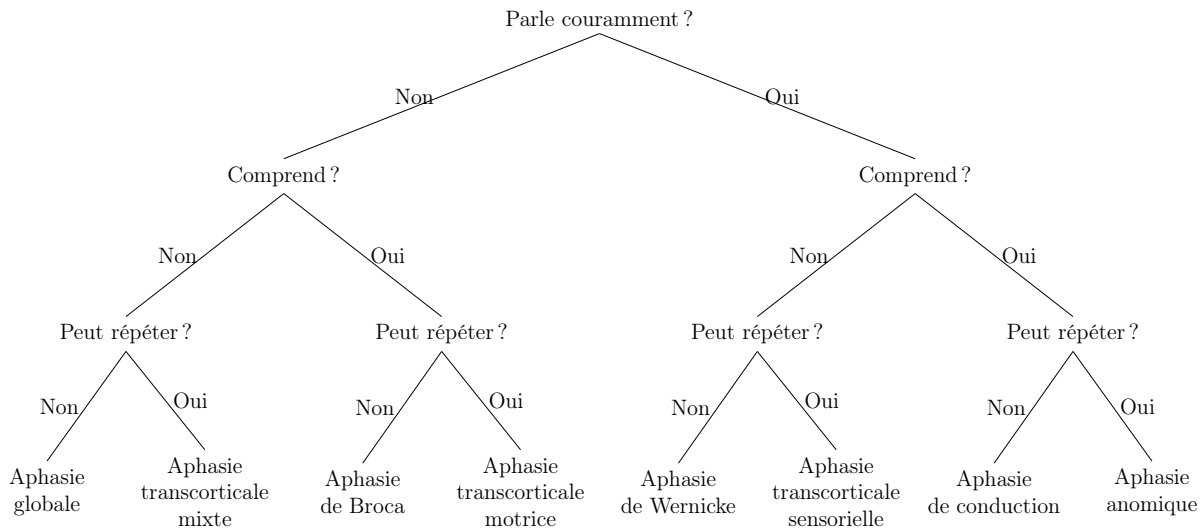


FIGURE 1.4 – Classification de certains types d'aphasie SREEDHARAN, 2018

Dans cette étude, nous nous intéressons principalement à l'aphasie de Broca. Il s'agit d'une aphasie expressive, c-à-d qui touche la capacité d'articuler sa pensée dans le langage et de le répéter, mais pas à celle de le comprendre (Voire la Figure 1.4).

1.4 Étiologie et épidémiologie de l'aphasie

Les AVC sont la première cause d'aphasie HALLOWELL, 2017. Ils représentent 75% des cas. Le traumatisme crânien en provoque 5% et les 20% restants se partagent entre les autres causes. L'âge est un facteur de risque très important pour les AVC, il l'est donc également pour l'aphasie. En effet, l'âge moyen des individus Français atteints de l'aphasie est 73 ans. 75% parmi eux sont âgés de plus de 65 ans dont 25% dépassent les 80 ans « DOSSIER DE PRESSE. Fédération Nationale des Aphasiques de France /12. » s. d.

Il est difficile d'estimer l'incidence et la prévalence globales de l'aphasie. Ceci est due au manque de données dans la majorité des pays du monde. Selon l'association nationale de l'aphasie « National Aphasia Association », s. d., 2 millions Américains en souffrent, soit 0.6%. En France, ce chiffre est de l'ordre de 300000 cas totaux et 30000 cas chaque année « DOSSIER DE PRESSE. Fédération Nationale des Aphasiques de France /12. » s. d. Ceci donne une prévalence de 0.44% et un taux d'incidence 0.044%.

33% des AVC résultent en une aphasie « DOSSIER DE PRESSE. Fédération Nationale des Aphasiques de France /12. » s. d. Un an après l'AVC, 13% des patients développent une aphasie de Broca.

Chapitre 2

Apprentissage séquence à séquence

Les modèles “séquence à séquence” sont une famille d’algorithmes de apprentissage automatique (ML) dont l’entrée et la sortie sont des séquences. Plusieurs tâches de apprentissage automatique, notamment en traitement automatique du langage (NLP), peuvent être formulées comme tâches d’apprentissage séquence à séquence. Parmi ces tâches, nous citons : la création de chatbots, la réponse aux questions, la reconnaissance automatique de la parole et la traduction automatique.

Dans ce chapitre, nous commençons par formuler le problème de modélisation de séquences. En suite, nous présentons les architectures neuronales les plus utilisées pour cette tâche. En fin, nous terminons avec une étude comparative de celles-ci.

2.1 Énoncé du problème

Formellement, le problème de modélisation séquence à séquence est celui de calculer une fonction partielle $f : X^* \rightarrow Y^*$, où :

- X est un ensemble dit d’entrées.
- Y est un ensemble dit de sorties.
- Pour un ensemble A , $A^* = \bigcup_{n \in \mathbb{N}} A^n$ est l’ensemble de suites de longueur finie d’éléments de A .

f prend donc un $x = (x_1, x_2, \dots, x_n) \in X^n$ et renvoie un $y = (y_1, y_2, \dots, y_m) \in Y^m$. Dans le cas général, $n \neq m$ et aucune hypothèse d’alignement n’est supposée. Il est souvent de prendre $X = \mathbb{R}^{d_e}$ et $Y = \mathbb{R}^{d_s}$ avec $d_e, d_s \in \mathbb{N}$. Dans ce cas, $x \in \mathbb{R}^{d_e n}$ et $y \in \mathbb{R}^{d_s m}$. Les indices peuvent représenter une succession temporelle ou un ordre plus abstrait (comme celui des mots dans une phrase).

La majorité des outils mathématiques historiquement utilisées pour ce problème viennent de la théorie du traitement de signal numérique. Cependant, l’approche actuellement dominante et celle qui a fait preuve de plus de succès, est de les combiner avec les réseaux de neurones.

2.2 Réseaux de neurones feed-forward

Les réseaux de neurones profonds sont parmi les modèles les plus expressifs en ML (apprentissage automatique). Leur succès pratique est incomparable aux modèles qui les ont précédés, que se soit en termes de qualité des résultats ou de variétés de domaines d'application. De plus, grâce aux théorèmes dits d'approximation universelle, ce succès empirique est formellement assuré.

2.2.1 Généralités

Dans cette section, nous introduisons les FFN, l'architecture neuronale la plus simple et la plus utilisée. Il s'agit d'une simple composition de couches affines avec des activations non affines (voire Définition 1).

Définition 1 (FFN).

Soient $k, w_0, w_1, \dots, w_{k+1} \in \mathbb{N}$, un réseau de neurones feed-forward de profondeur $k + 1$, à w_0 entrées et w_{k+1} sorties, est défini par une fonction :

$$\begin{cases} \mathbb{R}^{w_0} \rightarrow \mathbb{R}^{w_{k+1}} \\ x \mapsto \varphi_{k+1} \circ A_{k+1} \circ \varphi_k \circ A_k \circ \dots \circ \varphi_1 \circ A_1(x) \end{cases} \quad (2.1)$$

Où les A_i sont des fonctions affines $\mathbb{R}^{w_{i-1}} \rightarrow \mathbb{R}^{w_i}$ et les φ_i sont des fonctions quelconques, typiquement non affines $\mathbb{R}^{w_i} \rightarrow \mathbb{R}^{w_i}$, dites *d'activations*. La fonction $\varphi_i \circ A_i$ est appelée la i^{e} couche du réseau.

Un tel réseau de neurones est souvent représenté par un graphe orienté acyclique $(k+1)$ -partie appelé son "architecture sous-jacente" (KEARNS & VAZIRANI, 1994). La Figure 2.1 illustre l'architecture d'un FFN de profondeur 3 avec $(w_0, w_1, w_2, w_3) = (4, 5, 5, 3)$.

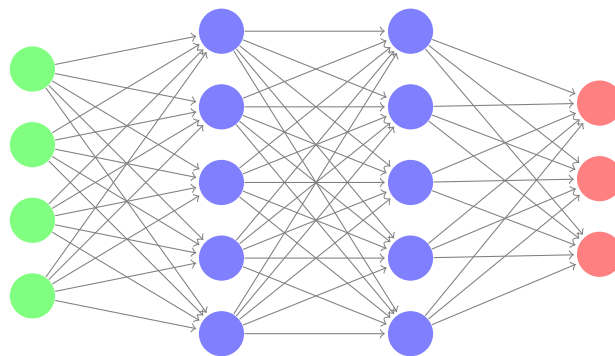


FIGURE 2.1 – Architecture sous-jacente d'un FFN de profondeur 3.

Deux FFN peuvent avoir la même architecture sous-jacente, en effet, cette dernière ne dépend que des dimensions de leurs couches respectives. De ce fait, une méthode de trouver pour une architecture et une fonction cible données le meilleur FFN est nécessaire.

Pour ce faire, nous exploitons le fait que les A_i soient des applications affines sur des espaces de dimensions finies. Nous pouvons donc les écrire comme combinaisons de

produits matriciels et de translations. Le problème se réduit donc à régler ¹ les paramètres des matrices en question. Cela nécessite une façon de quantifier la qualité d’approximation d’une fonction f par une autre \hat{f} . L’analyse fonctionnelle nous en donne plusieurs, les équations 2.2 et 2.3 sont deux exemples récurrents de fonctions dites *de perte*.

$$L_1(f, \hat{f}) = \int \|f - \hat{f}\| \quad (2.2)$$

$$L_2(f, \hat{f}) = \int \|f - \hat{f}\|^2 \quad (2.3)$$

Ayant fixé une fonction de perte L , l’entraînement revient à un problème d’optimisation. Dans le cas particulier où L est différentiable, l’algorithme du gradient peut être utilisé pour trouver un minimum local. Les gradients sont calculés en utilisant une méthode de dérivation automatique comme la rétro-propagation.

2.3 Réseaux de neurones récurrents

L’RNN (réseau de neurones récurrent) est une architecture de réseau de neurone conçue pour la modélisation des séquences. Elle se base sur l’idée de boucle de rétroaction pour capturer les dépendances temporelles.

1. En ML, le terme “entraîner” est plutôt utilisé.

Chapitre 3

Traduction automatique

La traduction automatique (MT) est une branche du traitement automatique du langage (NLP). Elle étudie l'utilisation des systèmes informatiques pour traduire le texte ou la parole d'une langue (appelée source) vers une autre (appelée cible) CHAN, 2015.

Dans ce chapitre, on introduit la traduction automatique du texte pour donner un point de références aux discussions des chapitres suivants. On y discute les plus répandues du domaine ainsi que les développements les plus récents qu'il a vécu.

3.1 Classification des méthodes de traduction automatique

La classification des méthodes de MT la plus citée à travers la littérature, repose sur les outils mathématiques de celles-ci. On distingue notamment trois familles de méthodes YANG et al., 2020 :

0. Des méthodes basées sur des connaissances linguistiques (règles de traduction).
0. Des méthodes basées sur les statistiques.
0. Des méthodes basées sur les réseaux de neurones.

On les appelle respectivement Traduction automatique à base de règle (RMBT), Traduction automatique statistiques (SMT), Traduction automatique neuronale (NMT). Les méthodes dans chacune de ces trois catégories peuvent être encore classifiées COSTA-JUSSÀ et al., 2016 ; YANG et al., 2020, ce qui donne lieu à la hiérarchie représentée par la Figure 3.1.

Le reste de ce chapitre sera organisé selon la structure donnée par la Figure 3.1. Cela nous donne la structure suivante

- Traduction automatique à base de règles
- Traduction automatique statistique
- Traduction automatique neuronale classique
- Traduction automatique à base de transformeurs

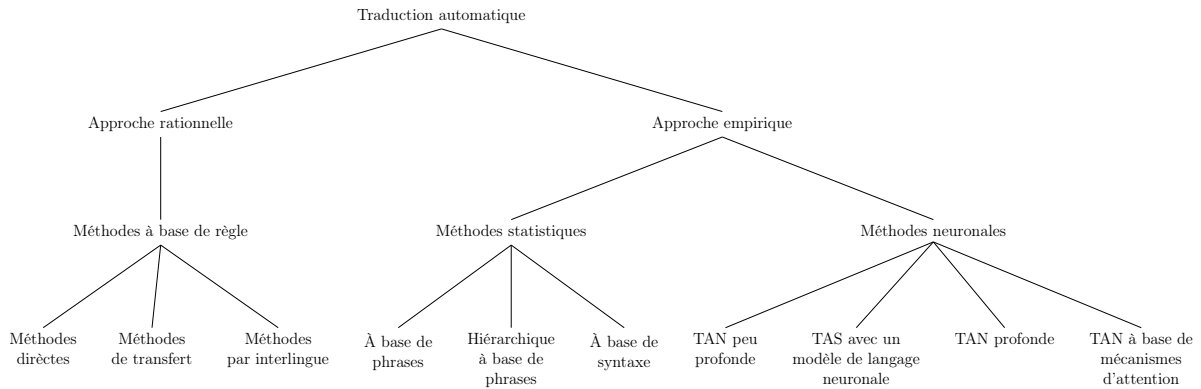


FIGURE 3.1 – Taxonomie des méthodes de traduction automatique.

3.2 Traduction automatique à base de règles

La traduction automatique à base de règle (RMBT) est historiquement le premier paradigme de MT. Étant apparue pendant les années 1950s, elle resterait l’approche dominante de MT jusqu’aux 1980s CHAN, 2015.

Comme son nom l’indique, la RMBT est basée sur des règles de traduction explicites, qui sont généralement créées manuellement à partir de connaissances linguistiques sur la langue source (LS) et la langage cible (LC). Les règles en question peuvent être d’ordre lexical (i.e des dictionnaires), syntaxique (i.e des grammaires) ou sémantique.

Toutes les méthodes de RMBT passent par deux phases : l’analyse de l’entrée dans la LS et la génération (ou synthèse) de la sortie dans la LC. Cependant, les règles utilisées dans ces deux phases peuvent varier en profondeur et dans le types de connaissances linguistiques employées. On distingue ainsi la RMBT en trois sous-familles de méthodes comme indiqué sur le sous-arbre gauche de la Figure 3.1.

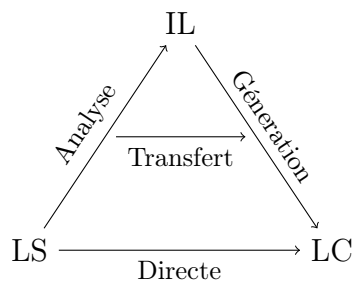


FIGURE 3.2 – Triangle de Vauquois

3.3 Traduction automatique statistique

Elle cherche à trouver des corrélations entre source et cible à partir de données parallèles. Elle requiert des pré-traitements comme alignement des mots et donne des Résultats

meilleurs que TABR.

3.4 Traduction automatique neuronale classique

Comme toutes les méthodes discutées dans ce chapitre, les méthodes neuronales passent par deux phases : celle d’analyse et celle de génération. Cela se fie naturellement au traitement par une architecture encodeur-décodeur.

3.5 Traduction automatique à base de transformeurs

Le transformeur (ou modèle auto-attentif) est un modèle de séquence proposé en 2017 VASWANI et al., 2017. Il vise à remédier aux lacunes des autres architectures utilisées en modélisation de séquences, notamment celles basées sur les RNN et les CNNs.

En effet, les RNN souffrent de deux problèmes fondamentaux. Le premier est celui du mémoire à court terme, c-à-d, la tendance à négliger les dépendances entre unités lexicales éloignées dans la séquence. Des variantes comme les LSTMs et surtout les GRUs ont réussi à éliminer ce problème au prix d’une augmentation dans la complexité des calculs. Cependant, ces variantes, ainsi que les RNN eux-mêmes, souffrent d’un deuxième défaut, celui du calcul séquentiel. Par conséquent, il n’est pas possible de paralléliser leurs calculs, ce qui les rends très lents VASWANI et al., 2017 ; YANG et al., 2020.

Les CNNs permet de contourner les inconvénients posés par les RNN. Ils sont entièrement parallélisables, ce qui les rends extrêmement rapides à dérouler sur GPU. De plus, ils peuvent calculer la dépendance entre deux positions arbitrairement éloignées. Or, le nombre d’opérations qu’ils utilisent pour ce faire augmente avec la distance entre les deux positions en question. De ce fait, ils sont également inadéquats pour la modélisation des corrélations à grande échelle YANG et al., 2020.

Le transformeur est une architecture à base de mécanisme d’attention qui offre une solution aux défis rencontrés par les autres architectures que nous avons introduites. Il ne souffre pas du mémoire à court terme et peut calculer les dépendances globales de la séquence en temps constant VASWANI et al., 2017. La Figure 3.3 illustre l’architecture d’un transformeur. Il s’agit d’une architecture encodeur-décodeur.

L’encodeur prend une séquence (x_1, x_2, \dots, x_n) de plongement de mots dans la LS et produit une suite (z_1, z_2, \dots, z_n) de représentations vectorielles (vecteurs de pensée). Cela est fait par le biais de l’Algorithme ?? ou.

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

Le décodeur prend cette suite de représentations et produit une troisième séquence (y_1, y_2, \dots, y_n) de mots dans LC.

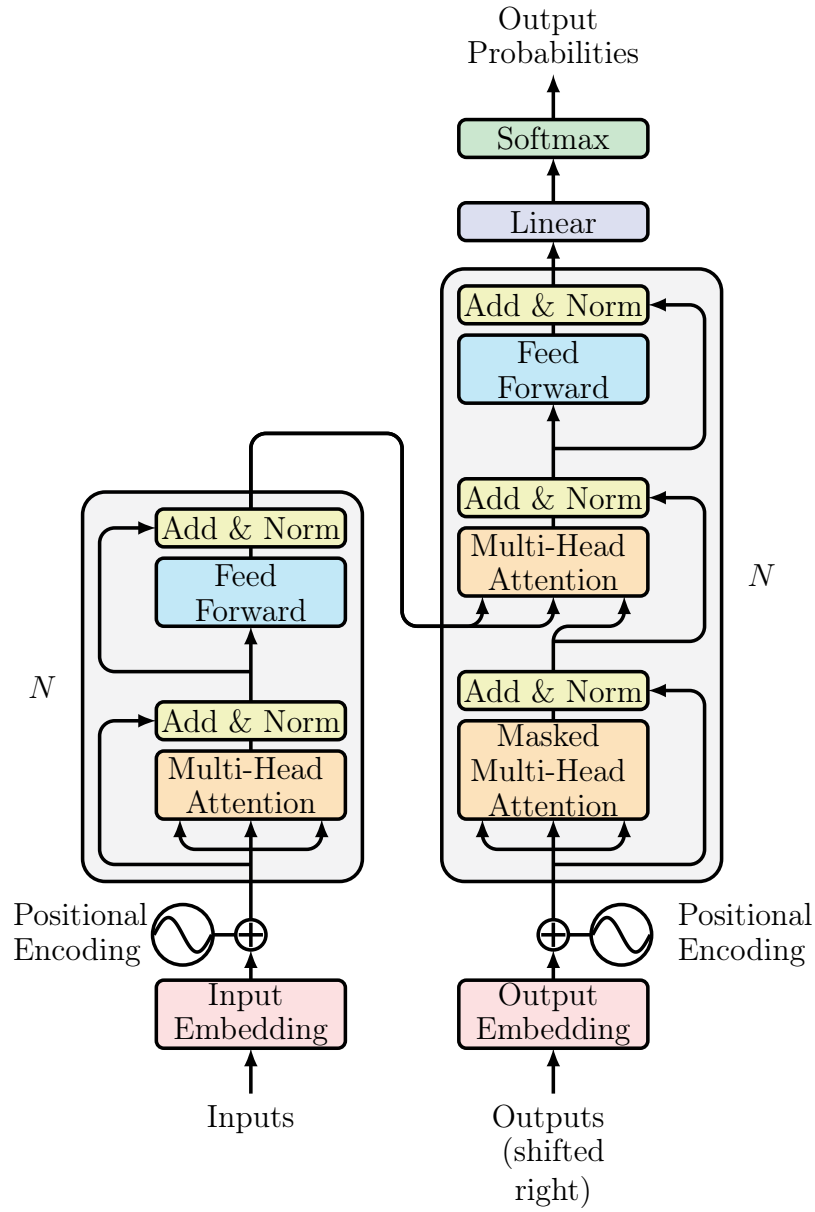


FIGURE 3.3 – L'architecture du transformeur VASWANI et al., 2017, Fig 1

Bibliographie

- BROCA, M. P. (s. d.). REMARQUES SUR LE SIÈGE DE LA FACULTÉ DU LANGAGE ARTICULÉ, SUIVIES D'UNE OBSERVATION D'APHÉMIE (PERTE DE LA PAROLE), 18.
- BRODMANN, K. (2007). *Brodmann's : Localisation in the Cerebral Cortex*. Springer Science & Business Media.
- CHAN, S.-w. (2015). *Routledge Encyclopedia of Translation Technology*. Routledge, Taylor & amp ; Francis Group.
- CHAPEY, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*. Wolters Kluwer Health/Lippincott Williams & Wilkins.
- COSTA-JUSSÀ, M. R., RAPP, R., LAMBERT, P., EBERLE, K., BANCHS, R. E., & BABYCH, B. (Éd.). (2016). *Hybrid Approaches to Machine Translation*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-21311-8>
- Dictionnaire de français Larousse. (s. d.). <https://www.larousse.fr/dictionnaires/francais/aphasie/4448>
- DOSSIER DE PRESSE. Fédération Nationale des Aphasiques de France /12. (s. d.). <https://docplayer.fr/229002408-Dossier-de-presse-federation-nationale-des-aphasiques-de-france-12-le-curtillard-le-haut-breda-telephone.html>
- FODOR, J. A. (1983). *The Modularity of Mind* [Google-Books-ID : 0vg0AwAAQBAJ]. MIT Press.
- HALLOWELL, B. (2017). *Aphasia and Other Acquired Neurogenic Language Disorders : A Guide for Clinical Excellence*. Plural Publishing.
- KEARNS, M. J., & VAZIRANI, U. (1994). *An Introduction to Computational Learning Theory* [Google-Books-ID : vCA01wY6iywC]. MIT Press.
- LORCH, M. (2011). Re-examining Paul Broca's initial presentation of M. Leborgne : Understanding the impetus for brain and language research. *Cortex*, 47(10), 1228-1235. <https://doi.org/10.1016/j.cortex.2011.06.022>
- MOHAMMED, N., NARAYAN, V., PATRA, D. P., & NANDA, A. (2018). Louis Victor Leborgne ("Tan"). *World Neurosurgery*, 114, 121-125. <https://doi.org/10.1016/j.wneu.2018.02.021>
- National Aphasia Association. (s. d.). <https://www.aphasia.org/>
- SREEDHARAN, S. (2018). *REAL-TIME FMRI BASED NEUROFEEDBACK FOR REHABILITATION OF POST-STROKE PATIENTS WITH APHASIA* (thèse de doct.). <https://doi.org/10.13140/RG.2.2.10868.37760/1>
- VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, & POLOSUKHIN, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>

YANG, S., WANG, Y., & CHU, X. (2020). A survey of deep learning techniques for neural machine translation. *arXiv preprint arXiv :2002.07526*.