

Mémoire de Master

Pour l'obtention du diplôme de Master 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

Résumé

L'aphasie est un trouble de langage qui résulte d'une lésion cérébrale (typiquement suite à un AVC). L'aphasie de Broca est une déficience de la production du langage causée par une lésion dans l'aire de Broca, une région du lobe frontal gauche du cerveau, responsable de la production de la parole. Une personne atteinte d'aphasie de Broca peut avoir des difficultés à articuler les mots et les phrases. Cependant, elle peut en général comprendre ce qui est dit. L'aphasie de Broca est associée à une diminution de la qualité de vie et à une augmentation du risque de dépression et de tentative de suicide.

La rééducation de la parole est le traitement le plus couramment prescrit aux personnes atteintes d'aphasie de Broca. En dépit de son efficacité, la rééducation de la parole est un traitement coûteux en termes de temps, argent et ressources humaines. Cela la rend indisponible à un grand nombre de personnes souffrant de l'aphasie de Broca.

L'utilisation des techniques basées sur le traitement automatique du langage pour améliorer la qualité de vie de ses individus est une voie d'exploration émergente qui a reçu beaucoup d'attention par les chercheurs pendant les années dernières.

Dans ce mémoire, nous nous intéressons à l'utilisation de la traduction automatique et la reconnaissance automatique de la parole pour automatiser une partie de la procédure de réhabilitation des personnes touchées par l'aphasie de Broca. Dans ce but, nous introduisons l'aphasie de Broca, ses causes, ses effets et les problèmes avec les traitements classiques. Ensuite, nous menons une étude bibliographique sur les travaux existant sur la traduction automatique et la reconnaissance automatique de la parole.

Mots clés — Aphasie de Broca, Apprentissage automatique, Traitement automatique du langage, Traduction automatique, Reconnaissance automatique de la parole.

Abstract

Aphasia is a language disorder caused by brain damage (most commonly a stroke). Broca's aphasia is a form of aphasia that impairs language production. It is caused by an injury to Broca's Area, an area of the frontal lobe of the brain; responsible for language decoding. A person suffering from Broca's aphasia may find it difficult to articulate words and sentences. However, they generally can understand what is said to them. This form of aphasia is associated with a lower quality of life and a higher risk of depression and suicide.

Speech therapy is the most commonly prescribed remedy to people with Broca's aphasia. Despite its effectiveness, it remains an expensive, time-consuming, and effort-heavy process. This makes it inaccessible to a significant number of people with aphasia.

The use of natural language processing-based techniques to improve these people's quality of life is an emerging research avenue that has enjoyed the attention of many researchers in recent years.

In this thesis, we are interested in the use of machine translation and automatic speech recognition to partially automate the rehabilitation of people with aphasia. To this end, we introduce aphasia, its causes, consequences, and the problems of classical treatment methods. We then undertake a bibliographical study of the existing works pertaining to machine translation and automatic speech recognition.

Keywords — Broca aphasia, Machine learning, Natural language processing, Machine translation, Automatic speech recognition.

ملخص

الحبسة إضطرابٌ لغوي ناتج عن تلف في الدماغ، غالبا نتيجة سكتة دماغية. حبسة بروكا حبسة تنتج عن إصابة في منطقة بروكا، وهي منطقة في الفص الجبهي الأيسر للدماغ تعنى بإنتاج الكلام. قد يجد المصاب بحبسة بروكا صعوبة في تكوين الجمل والكلمات، إلا أنه عادة يفهم ما يقال. ترتبط هذه الحبسة بتدني مسوى العيش وارتفاع خطر الاكتئاب والانتحار.

علاج النطق هو أكثر العلاجات وصفا للمصابين بحبسة بروكا. رغم نجاعته، فهو يظل مكلفا للوقت والمال والجهد، ما يحول دون توفره لعدد كبير ممن يحتاجونه.

توظيف تقنيات معالجة اللغة الطبيعية لتحسين حياة المصابين بحبسة بروكا مجال بحث حظي باهتمام العديد من الباحثين في الأعوام الأخيرة.

في هذه الأطروحة، نهتم باستعمال الترجمة الآلية والتعرف الآلي على الكلام لتأدية جزء من علاج النطق لحبسة بروكا أوتوماتيكيا. من أجل ذلك، نبدأ بالتعريف بحبسة بروكا أسبابا ونتائج، ثم نتطرق لعيوب العلاجات المعتادة. بعدها نعرض دراسة ببليوغرافية للأعمال التي سبق إنجازها في مجالي الترجمة الآلية والتعرف الآلي على الكلام.

الكلمات المفتاحية – حبسة بروكا، تعلم الآلة، معالجة اللغة الطبيعية، ترجمة آلية، تعرف آلي على الكلام.

Table des matières

Page de garde	i
Table des matières	iv
Table des figures	v
Liste des algorithmes	vi
Sigles et abréviations	vii
Introduction générale	1
I Contribution	3
1 Conception	4
1.1 Architecture générale de la solution	4
1.2 Reconnaissance automatique de la parole	6
1.2.1 Préparation des données	6
Conclusion générale	8
Bibliographie	9

Table des figures

1.1	Architecture générale de la solution.	5
-----	---	---

Liste des algorithmes

Sigles et abréviations

ASR	reconnaissance automatique de la parole
AVC	accident vasculaire cérébral
DL	apprentissage profond
ML	apprentissage automatique
MT	traduction automatique
NLP	traitement automatique du langage
NMT	traduction automatique neuronale
S2S	séquence-à-séquence

Introduction générale

L'aphasie est un trouble de communication qui complique un grand pourcentage d'accidents vasculaires cérébraux, une condition médicale qui touche plus de 12.2 millions de personnes par an dans le monde. Ce chiffre est susceptible d'augmenter avec l'augmentation de l'espérance de vie (FEIGIN et al., 2022).

L'aphasie de Broca est une forme d'aphasie qui affecte la capacité de s'exprimer oralement ou par écrit. Elle résulte d'une lésion dans l'aire de Broca, une région du cerveau qui est responsable de la production de la parole. Les personnes qui souffrent d'une aphasie de Broca ont des difficultés à produire des mots, mais peuvent comprendre ce qui est dit (CHAPEY, 2008).

Ces difficultés peuvent avoir des conséquences néfastes sur plusieurs aspects de la vie quotidienne. Ceci peut inclure la communication avec les proches, la participation à des activités sociales, l'exercice d'un emploi ou même la demande d'aide en cas d'urgence (HALLOWELL, 2017).

L'utilisation de techniques d'apprentissage automatique et du traitement automatique du langage au bénéfice des personnes atteintes de l'aphasie est une piste de recherche qui commence à capturer l'attention de plusieurs chercheurs (MISRA et al., 2022 ; QIN et al., 2022 ; SMAÏLI et al., 2022).

Dans ce mémoire, notre objectif est de fournir une revue de l'état de l'art sur les travaux qui ont été faits dans cette direction. Nous portons une attention particulière à la traduction automatique et à la reconnaissance automatique de la parole appliquées à l'aphasie de Broca.

Pour ce faire, nous avons organisé notre travail en trois chapitres :

1. ??.

Dans ce chapitre, nous présentons en général les domaines de recherche qui nous intéressent. À cette fin, le chapitre est divisé en trois sections :

- (1) La première section présente l'aphasie de Broca,
- (2) la deuxième introduit la traduction automatique
- (3) et la troisième présente la reconnaissance automatique de la parole.

2. ??.

Ce chapitre sert à familiariser le lecteur avec le cadre d'étude général dans lequel s'inscrivent la traduction automatique, la reconnaissance automatique de la parole

et la majorité des tâches de traitement automatique du langage. Il s'agit de la modélisation de séquences. Nous y présentons l'énoncé du problème et les différentes architectures neuronales qui ont été proposées pour le résoudre.

3. ??.

Ce chapitre part de l'étude générale faite dans le chapitre ??. Il détaille l'application des meilleures architectures neuronales qui y sont présentées dans le cadre de la traduction automatique et de la reconnaissance automatique de la parole.

Première partie

Contribution

Chapitre 1

Conception

Dans les chapitres précédents, nous avons effectué une étude bibliographique sur l'aphasie de Broca et les méthodes de traitement automatique du langage (NLP, de l'anglais : natural language processing) qui peuvent être utilisées pour la traiter (particulièrement les modèles séquence-à-séquences (S2S)). Cela nous a permis de développer une idée claire d'un système S2S pour la réhabilitation de la parole aphasique.

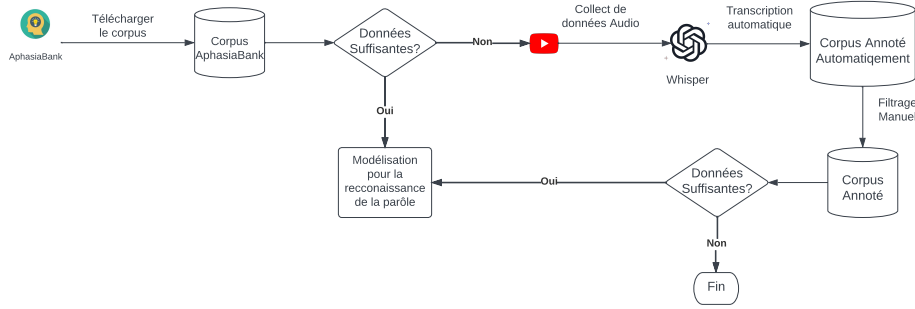
Dans ce chapitre, nous allons présenter les détails de la conception de notre système. Nous commençons par décrire la démarche suivie pour le concevoir. Puis, nous présentons l'architecture générale du système, que nous détaillons par la suite.

1.1 Architecture générale de la solution

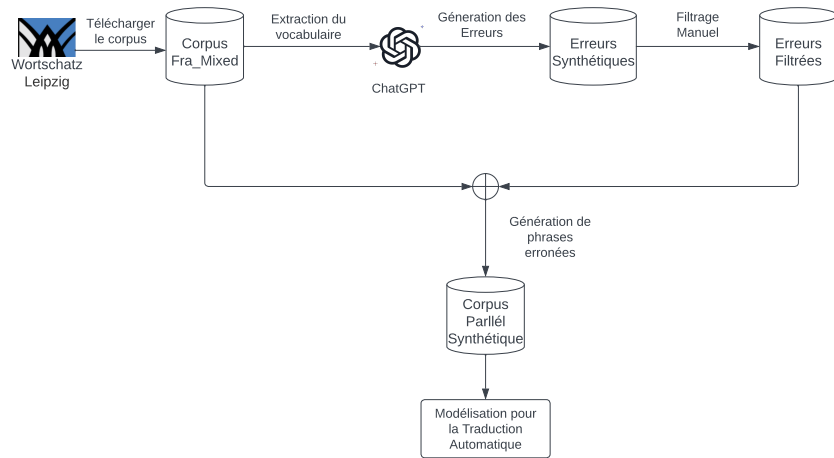
Pour résoudre le problème de la réhabilitation de la parole aphasique, nous proposons un système dont l'architecture générale est illustrée dans la figure 1.1. Ce système est composé de deux parties principales (a) le sous-système d reconnaissance automatique de la parole (ASR, de l'anglais : automatic speech recognition) qui permet de transcrire la parole aphasique en texte et dont l'architecture est illustrée dans la Figure 1.1a. (b) le sous-système de traduction automatique neuronale (NMT, de l'anglais : neural machine translation) qui permet de traduire le texte transcrit en parole saine et dont l'architecture est illustrée dans la Figure 1.1b.

Pour la partie ASR, nous avons choisi de l'organiser en deux activités principales :

- (i) Préparation d'un corpus de donnée (parole/écrit), ce qui passe par :
 - La collecte de données existantes : consultation de corpus existants et de bases de données de parole aphasique, téléchargement et évaluation de ces données. Si les données collectées sont suffisantes (en qualité et en volume), passer à la modélisation, sinon, passer à l'étape suivante.
 - La collecte de données supplémentaires : repérage et téléchargement de vidéos de parole aphasique sur internet.
 - Transcription des données collectées : dans un premier temps, les vidéos sont transcrites automatiquement à l'aide d'un modèle ASR existant.



(a) Déroulement de la partie ASR.



(b) Déroulement de la partie NMT.

FIGURE 1.1 – Architecture générale de la solution.

- Filtrage manuel : les transcriptions automatiques sont filtrées manuellement pour corriger les erreurs de transcription. Le résultat de cette étape est un corpus de vidéos transcrites semi-automatiquement.
- Évaluation du corpus : si le corpus est suffisant (en qualité et en volume), passer à la modélisation, sinon, passer à la partie NMT.

(ii) Création et entraînement d’un modèle sur ce corpus.

Une organisation similaire est adoptée pour la partie NMT :

- (i) Création d’un corpus parallèle (parole aphasique/parole saine) : contrairement à la partie ASR, aucun corpus parallèle n’est disponible pour l’aphasie de Broca. Dans l’absence de possibilité d’en créer un, nous avons choisi de créer un corpus parallèle synthétique. Cela peut être réalisé en suivant la procédure suivante :
 - Prendre un corpus du Français.
 - Extraire le vocabulaire de ce corpus.
 - Sélection des mots : choisir dans ce vocabulaire les mots qui sont “*fréquents*” “*difficiles*” à prononcer.

- Création des erreurs : passer ces mots au modèle chatGPT pour générer des variantes erronées.
 - Filtrage manuel : traiter manuellement les erreurs générées pour ne garder que celles qui sont similaires aux erreurs produites dans l’aphasie de Broca.
 - Création du corpus parallèle : à partir du corpus du original et des erreurs filtrées, créer un corpus parallèle.
- (ii) Création et entraînement d’un modèle sur ce corpus : une fois le corpus parallèle créé, il est possible de créer un modèle de traduction et de l’entraîner sur ce corpus. La démarche traditionnelle de apprentissage profond (DL, de l’anglais : deep learning) est suivie pour cette étape :
- Division du corpus : le corpus est divisé en trois parties : 1. corpus d’entraînement, utilisé pour entraîner le modèle, 2. corpus de validation, utilisé pour évaluer le modèle pendant l’entraînement et 3. corpus de test, utilisé pour évaluer le modèle après l’entraînement.
 - Création du modèle : dans cette étape, l’architecture du modèle est choisie en fonction de la tâche en question.
 - Entraînement du modèle : le modèle est entraîné sur le corpus d’entraînement pour ajuster ses paramètres.
 - Évaluation du modèle : le corpus de test est utilisé pour mesurer la performance du modèle. Plusieurs métriques peuvent être utilisées pour cette évaluation. Si les résultats sont satisfaisants, passer à l’étape suivante, sinon, passer à l’étape précédente.

Dans la suite de ce chapitre, nous reprenons dans le détail les différentes étapes de cette architecture, que nous avons abordées brièvement dans cette section.

1.2 Reconnaissance automatique de la parole

Le modèle d’ASR permet de transformer la parole aphasique en texte dans le but de l’utiliser comme entrée pour le modèle de traduction. Les étapes de sa construction sont établies dans la section précédente. Dans cette section, nous allons détailler ces étapes.

1.2.1 Préparation des données

La première étape de tout projet de apprentissage automatique (ML, de l’anglais : machine learning) est la préparation des données. Dans ce cas, le jeu de données doit être constitué de couples de morceaux d’audio et de leurs transcriptions textuelles.

Choix du corpus et collecte de données existantes

Notre choix s’est porté sur le corpus *AphasiaBank* (MACWHINNEY et al., 2011). Il fait partie du projet *TalkBank* (MACWHINNEY, 2007), une collection de bases de données

créées pour l'étude du langage. AphasiaBank contient plusieurs vidéos d'entre-entrevues entre des chercheurs et des personnes souffrant de l'aphasie de Broca. Ses vidéos sont accompagnées par des transcriptions textuelles faites par des experts dans un format particulier. La qualité des transcriptions est donc excellente.

Cependant, le volume de données sur l'aphasie de Broca est très limité. En effet, un seul des 11 exemples disponibles en Français est un exemple de l'aphasie de Broca. Il s'agit d'une vidéo de 12 min 03 s, qui contient 3000 mots. Il faut noter que la moitié de ces mots sont prononcés par le chercheur. La durée effective de la parole aphasique est donc de 6 min. Cela est de très loin insuffisant pour entraîner un modèle profond. Pour ce but, il est nécessaire de collecter des données supplémentaires.

Collecte de données supplémentaires

Les données d'AphasiaBank étant insuffisantes, d'autres sources sont nécessaires. Plusieurs enregistrements de personnes souffrant de l'aphasie de Broca sont disponibles sur internet (YouTube, Vimeo, ...). La qualité de ces enregistrements est très variable et largement inférieure à celle d'AphasiaBank. Cependant, leur ajout au corpus est nécessaire pour augmenter sa taille.

Notre recherche nous a permis d'obtenir 22 enregistrements de personnes souffrant de l'aphasie de Broca d'une durée totale de 48 min 44 s. Parmi ses enregistrements, 7 sont des hommes et 13 sont des femmes et 2 sont des enregistrements de groupes.

Transcription et filtrage des données

Les 22 enregistrements collectés sont transcrits à l'aide de Whisper (RADFORD et al., 2022). Cela permet d'avoir des transcriptions textuelles de qualité. Cependant, les transcriptions obtenues contiennent des erreurs (particulièrement pour les prononciations aphasiques).

L'étape suivante est donc de filtrer à la main les transcriptions obtenues. Cela permet de corriger les erreurs de transcription et de réintroduire les prononciations aphasiques éliminées par Whisper. La sortie de cette étape est un corpus (parole/écrit).

Les données d'AphasiaBank sont déjà transcrits, mais cela est fait dans un format particulier. Il est donc nécessaire de réécrire les transcriptions en Français standard. L'exemple d'AphasiaBank est donc ajouté au corpus, ce qui fait un total de 1 h 0 min 47 s de parole aphasique. Cela est encore insuffisant pour entraîner un modèle profond. Cependant, il peut servir aux chercheurs qui souhaitent travailler sur l'aphasie de Broca. On rend donc disponible ce corpus.

Conclusion générale

L’aphasie de Broca est un trouble de communication qui touche une partie grandissante de la population mondiale. Faisant souvent suite à un accident vasculaire cérébral (AVC), l’aphasie de Broca affecte le décodage des mots et la production de la parole. Elle diminue ainsi mesurablement la qualité de vie des personnes qui en sont atteintes (CHAPEY, 2008 ; FEIGIN et al., 2022 ; ROSS & WERTZ, 2010).

Le traitement le plus courant de l’aphasie de Broca est la rééducation de la parole par un orthophoniste. En dépit d’être un traitement efficace, il est gourmand en temps, en argent et en ressources humaines, ce qui en fait un traitement peu accessible pour la majorité des personnes atteintes. Ce manque d’accessibilité, combiné à la gravité de certaines conséquences de l’aphasie de Broca, rend urgent le développement de solutions alternatives (da FONTOURA et al., 2012 ; FLOWERS et al., 2016).

Les méthodes informatiques, notamment les techniques de ML et de NLP, semblent avoir le potentiel de réduire les coûts matériels et humains associés à la rééducation de la parole. Elles peuvent ainsi faciliter l’accès au traitement de l’aphasie de Broca (MISRA et al., 2022 ; QIN et al., 2022 ; SMAÏLI et al., 2022).

Dans ce mémoire, nous avons exploré la possibilité d’utiliser la traduction automatique (MT, de l’anglais : machine translation) et l’ASR, deux techniques de NLP basées sur l’apprentissage S2S pour aider les personnes atteintes de l’aphasie de Broca.

Nous avons d’abord introduit l’aphasie pour familiariser le lecteur avec ses causes, sa portée, ses effets, les traitements disponibles et les défis auxquels ils sont confrontés. Après cela, nous avons présenté la modélisation S2S, le cadre général de la MT et de l’ASR. Nous avons posé le problème et présenté les solutions que nous avons évaluées et comparées. Le résultat de cette comparaison est une supériorité nette du transformeur sur les autres modèles. Nous avons alors exploré dans le troisième chapitre les différentes publications qui ont étudié l’utilisation du transformeur pour la résolution de problèmes de MT et d’ASR.

Bibliographie

- CHAPEY, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders*. Wolters Kluwer Health/Lippincott Williams & Wilkins.
- da FONTOURA, D. R., RODRIGUES, J. d. C., CARNEIRO, L. B. d. S., MONÇÃO, A. M., & de SALLES, J. F. (2012). Rehabilitation of language in expressive aphasia : a literature review. *Dementia & Neuropsychologia*, 6(4), 223-235. <https://doi.org/10.1590/S1980-57642012DN06040006>
- FEIGIN, V. L., BRAININ, M., NORRVING, B., MARTINS, S., SACCO, R. L., HACKE, W., FISHER, M., PANDIAN, J., & LINDSAY, P. (2022). World Stroke Organization (WSO) : Global Stroke Fact Sheet 2022. *International Journal of Stroke : Official Journal of the International Stroke Society*, 17(1), 18-29. <https://doi.org/10.1177/17474930211065917>
- FLOWERS, H., SKORETZ, S., SILVER, F., ROCHON, E., FANG, J., FLAMAND-ROZE, C., & MARTINO, R. (2016). Poststroke Aphasia Frequency, Recovery, and Outcomes : A Systematic Review and Meta-Analysis. *Archives of Physical Medicine and Rehabilitation*, 97, 2188-2201. <https://doi.org/10.1016/j.apmr.2016.03.006>
- HALLOWELL, B. (2017). *Aphasia and Other Acquired Neurogenic Language Disorders : A Guide for Clinical Excellence*. Plural Publishing.
- MACWHINNEY, B. (2007). The talkbank project. *Creating and Digitizing Language Corpora : Volume 1 : Synchronic Databases*, 163-180.
- MACWHINNEY, B., FROMM, D., FORBES, M., & HOLLAND, A. (2011). AphasiaBank : Methods for studying discourse. *Aphasiology*, 25(11), 1286-1307. <https://doi.org/10.1080/02687038.2011.589893>
- MISRA, R., MISHRA, S. S., & GANDHI, T. K. (2022). Assistive Completion of Agrammatic Aphasic Sentences : A Transfer Learning Approach using Neurolinguistics-based Synthetic Dataset [arXiv :2211.05557 [cs, q-bio]], (arXiv :2211.05557). <http://arxiv.org/abs/2211.05557>
- QIN, Y., LEE, T., KONG, A. P. H., & LIN, F. (2022). Aphasia Detection for Cantonese-Speaking and Mandarin-Speaking Patients Using Pre-Trained Language Models. *2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, 359-363. <https://doi.org/10.1109/ISCSLP57327.2022.10037929>
- RADFORD, A., KIM, J. W., XU, T., BROCKMAN, G., MCLEAVEY, C., & SUTSKEVER, I. (2022). Robust Speech Recognition via Large-Scale Weak Supervision [arXiv :2212.04356 [cs, eess]], (arXiv :2212.04356). <http://arxiv.org/abs/2212.04356>
- ROSS, K., & WERTZ, R. (2010). Quality of life with and without aphasia. *Aphasiology*. <https://doi.org/10.1080/02687030244000716>

SMAÏLI, K., LANGLOIS, D., & PRIBIL, P. (2022). Language rehabilitation of people with BROCA aphasia using deep neural machine translation. *Fifth International Conference Computational Linguistics in Bulgaria*, 162.