

Troubles du langage et de la pensée : campagne d'annotation

Alexis Biver - Emile Colin - Corentin Lefebvre

CONTEXTE

Ce projet s'inscrit dans le projet de recherche Schizophrénie et Langage : Analyse et Modélisation (SLAM) porté par la Maison des Sciences de l'Homme (MSH).

Il part de l'hypothèse que les singularités dans le discours d'un schizophrène ne seraient pas sémantiques mais pragmatiques et dues à une mauvaise utilisation des règles sous-jacentes du dialogue. En effet, les experts ont mis en lumière la présence de ruptures dans des dialogues avec un schizophrène.

Il était donc nécessaire de confirmer les observations des chercheurs par le biais d'expérimentations, et cela via l'annotation de textes comportant ces ruptures par des annotateurs "naïfs", qui ne cherchent donc pas ce phénomène.

CAMPAGNE D'ANNOTATION

Nous avons repris le protocole de la campagne précédente en y ajoutant deux textes témoins et un nouveau texte à rupture discursive. Nous avons effectué un recrutement parmi nos proches et quelques volontaires pour des passations d'une durée moyenne de 1h. Parmi 6 retranscriptions de dialogues entre un psychologue et un patient, chaque volontaire a annoté 3 textes (un témoin et deux à rupture) sélectionnés au hasard.

Nous avons finalement réussi à obtenir un panel de 38 annotateurs variés et représentatifs ainsi que des retours intéressants via un questionnaire.

ANALYSE DES DONNÉES RECUEILLIES

La plus grande partie de notre travail a concerné l'analyse des données issues de cette campagne. Nous avons effectué nos analyses en Python et présenté nos résultats dans un notebook Jupyter pour que nos scripts puissent être utilisés par des non-informaticiens.

Nous avons étudié d'une part les différences dans la manière d'annoter entre les textes témoins et les textes à rupture, et d'autre part le regroupement des annotateurs selon la proximité de leurs annotations. Nous n'avons pas trouvé significativement plus d'annotations présentant de rupture de la frontière droite dans les textes à ruptures que dans les textes témoins, mais il semble que l'annotation des premiers soit moins consensuelle que celle des seconds. En outre, l'annotation des interventions du psychologue fait plus consensus que celles des interventions du patient.

De plus par regroupements nous avons pu identifier dans 2 textes à ruptures une forte proximité entre une majorité d'annotateurs et un expert.