

## SADSUI

### Sociolinguistique et Analyse de Discours de personnes SUlcidaires : analyse des conditions de (dys-)félicité d'un dispositif chat de prévention

Guðrun LEDEGEN, responsable scientifique, Professeure des Universités Sciences du Langage  
(Sociolinguistique – Analyse de discours)

Le projet SADSui (Sociolinguistique et Analyse de Discours de personnes SUlcidaires) est un projet d'analyse interdisciplinaire d'un dispositif numérique de prévention du suicide par chat par une analyse de discours outillée par la lexicométrie. Il réunit des enseignants-chercheurs venant des sciences du langage (sociolinguistique, analyse de discours et traitement automatique du langage), de la psychologie (psychologie sociale, psychologie du travail, ergonomie) et des sciences de l'informatique.

Il explore un grand corpus de chat (2005-2015), et cette analyse est complétée d'enquêtes et d'entretiens compréhensifs auprès des utilisateurs et des bénévoles du dispositif numérique de prévention.

Le travail analytique du corpus, interdisciplinaire, combine les analyses manuelles et automatisées, réunissant ainsi les approches quantitatives (mathématique et statistique) et qualitatives (descriptive et interprétative), afin d'analyser les conditions de *félicité* (réussite) et de *dysfélicité* (moindre réussite) dans ces interactions particulières, où l'empathie, portée par la voix humaine dans l'interaction téléphonique par exemple, doit s'exprimer avec d'autres moyens que l'intonation et la qualité de la voix.

Les buts des analyses du corpus sont de mieux comprendre le fonctionnement de ces conversations sensibles et, à partir de ces données réelles, de compléter les formations des bénévoles écoutants de l'association de prévention. Ces buts d'amélioration du dispositif d'écoute seront aussi atteints grâce à des entretiens compréhensifs et des enquêtes auprès des bénévoles écoutants, centrés sur l'interface chat, par un retour sur leurs pratiques personnelles et sur nos analyses qui leur sont présentées, ainsi qu'auprès des appelants.

Le travail qui a pu être effectué grâce au programme *Marsouin* porte sur la première phase d'analyse, axée sur le corpus de données en soi :

- Un premier travail de mise aux normes du RGPD (règlement général de protection des données) du corpus a demandé à **anonymiser** toutes les informations personnelles, i.e. toutes les entités nommées (nom, prénom, lieu, institution, ...) en les remplaçant par des codes, tout en gardant le cas échéant le lien entre différentes conversations, pour les rares personnes qui appellent à plusieurs reprises (poly-appelants, cf. plus bas). En effet, tous les chercheurs du projet n'accèdent qu'aux données anonymisées, et ils s'engagent par ailleurs à ne jamais publier d'extraits longs des conversations étudiées, mais uniquement de courts extraits non rattachables à une personne en particulier, et les résultats globaux extrapolables aux différentes catégories étudiées (cf. ci-dessous).
- Ensuite, pour permettre l'interrogation du corpus, ce dernier est **balisé** par des **codes** spécifiques à chaque interface logiciel : nous utilisons ainsi les logiciels Lexico 3 et Lexico 5, Hyperbase, TXM et Tropes.
- Enfin, dans chaque format particulier sont créées des **catégories de métadonnées** afin de pouvoir comparer les différentes conversations, selon les **dates** (jour, mois, année) ou les **heures** de connexion, selon la **durée** ou le **rythme** de saisie (les *time-codes* précédant chaque prise de parole dans le corpus, permettant des calculs fins de rythme moyen pour chaque conversation, mais aussi des accélérations ou ralentissements, dont les silences, durant les conversations).
- D'autres métadonnées sont créées à partir du corpus même, tels que le **genre** des appelants, leur **âge**, leur **métier** ou leurs **études**, leur **raison** de venue sur le site ... Cette deuxième série de métadonnées a été réalisée **manuellement puis semi-automatiquement**, en prenant appui sur des structures menant de façon indéniable<sup>1</sup> à la définition du genre de l'appelant (*je* + verbes d'état + formes adjectivales ou nominales audibles (ex. : *petit/petite ; une femme/un homme ; ...*) ou son âge (*je* + *avoir* + nombre + *âge* ; ...). L'identification du métier ou du niveau d'études n'a pu encore aboutir totalement, ainsi que la raison de venue sur le site (souvent de multiples raisons sont mentionnées avant que l'appelant n'en vienne à la raison principale : cette étude mérite un approfondissement qualitatif d'envergure).

---

<sup>1</sup> Ainsi, nous ne faisons confiance aux catégories des très grandes bases de données, comme Google, surtout dans le contexte sensible de nos données, où le genre peut être une difficulté personnelle pour les appelants.

Une fois, ces différents éléments mis à disposition, nous avons pu commencer à explorer différents aspects du corpus : ainsi,

- les **marqueurs discursifs** qui caractérisent pleinement le genre discursif identifié : différentes pratiques distanciantes, dont par exemple le *nous* 'de distance' (utilisé avec des verbes d'action ou d'émotion plutôt portés par l'unique personne en ligne : *nous ressentons votre tristesse*), permettent aux écoutants de se mettre à distance de l'exercice d'écoute mené, et de se préserver ainsi de verser trop de sympathie, en restant à un niveau d'empathie. Cette pratique méthodologique présente une grande pertinence autant pour les écoutants que pour les appelants, qui viennent à l'écran en question justement pour être en interaction avec une personne inconnue, qui n'est pas impliquée dans leur situation vécue ; elle est souvent imaginée plus âgée, et constitue une ressource à qui se confier et dont recevoir un retour éclairant.
- l'**émotion dite** désigne l'expression de façon claire et constante des émotions (et non de façon contextuelle : par exemple, le terme *ami* ne porte pas un sème d'émotion de façon constante, mais bien dans le cas de la conversation portant sur la perte dudit ami)). Cette émotion dite est explorée dans le corpus : quels termes sont utilisés, est-ce que les termes d'émotion utilisés par les écoutants sont des répétitions des termes de l'appelant ou des expressions personnelles, à quels moments des interactions interviennent-ils ? ... Une exploration détaillée qualitative de plusieurs conversations nous a ainsi permis par ailleurs de voir cette émotion dite à l'œuvre au fil du rythme de la conversation : ainsi, il est intéressant de noter que la conversation commence souvent avec des émotions *dysphoriques* (moins heureuses) pour évoluer vers des émotions davantage *euphoriques* (heureuses) en fin de conversation. Ce travail est effectué par l'écoutant qui réussit l'identification d'éléments positifs dans le discours de l'appelant, et ce changement de tonalité apporte des temps de compliment, d'humour ..., contexte qui permet souvent à l'appelant de livrer la confiance attendue (ce pour quoi le moment est difficile pour elle / lui).

En plus, il est intéressant de noter que le rythme de la conversation s'amplifie très fortement dans ces moments d'accroissement émotionnel, donnant lieu à ce qu'on peut appeler le chevauchement des tours de frappe. L'inscription de ces temps forts d'émotions dans le déroulement des conversations – à l'intérieur des différentes phases entamées dans l'exposé narratif ou argumentatif, ... – nous apparaissent souvent coïncider avec ces moments de la confiance.

- Une première exploration de **typologisation des écoutants** a pu s'effectuer sur la base de l'examen des différents écoutants en interaction avec deux poly-

appelants, phénomène très rare dans le corpus : l'étude portait sur une personne adulte, et un enfant, qui multiplie sur plusieurs mois les appels ; une remarquable distinction entre les façons de s'adresser à l'adulte et au jeune s'est ainsi fait jour : si l'enfant est d'une part conforté dans certaines de ses décisions fort concrètes, il se voit aussi adresser un très grand nombre de généralités (dictons ('le soleil se lève toutefois le matin'), tautologies ('un être qui souffre est un être qui souffre'), ...). Cette pratique positionne l'écouter d'autant plus comme un adulte qui conseille du fait de son expérience et de sa sagesse. L'adulte en revanche va davantage recevoir des demandes d'explicitation, et des contre-argumentations, les interlocuteurs étant davantage hiérarchiquement équivalents.

Les explorations du corpus futures porteront aussi sur la gestion et l'évolution des **silences** (par quoi remplacer un hm hm d'acquiescement à l'écran ?) et de la **punctuation** : il s'avère en effet que sur l'étendue de 11 ans, une évolution de la punctuation se soit faite jour, par l'émergence de la disparition de la punctuation finale par le point à la dernière phrase saisie. Plusieurs études font état d'une évaluation négative (très forte insistance, agressivité) en cas de présence de la punctuation : ainsi « OK. » est ressenti comme une expression de mauvaise humeur ou de défi, par les jeunes. De la même façon, la valeur des points de suspensions (...) semblant impliquer des interprétations avec beaucoup d'implicite, de nouvelles formes à double point ont vu le jour (..), autant chez les appelants que chez les écoutants. Leurs valeurs et évolutions sont à l'étude actuellement.

Il reste des explorations innombrables à mener sur ces conversations : elles permettront

- d'éclairer les interactions sur les chats de conversation en général, et dans le cadre professionnel distant en particulier,
- d'identifier une pratique professionnelle d'accueil de la parole sensible et de la personne en détresse, en déployant une empathie uniquement par cette voie de l'écrit,
- de cerner les pratiques et attentes des appelants, en particulier dans le cadre de ce chat, par comparaison avec les autres moyens mis à leur disposition (mail et téléphone) (cf. les entretiens compréhensifs à venir),
- ...