

Sémantique inférentielle et compréhension des verbatim clients

Geoffrey Doucy
viavoo.

Taoufik Massoussi
LDI, UMR 7187, Université Paris 13-CNRS – viavoo.

Introduction

L'un des enjeux du Traitement Automatique des Langues est l'analyse et l'étiquetage sémantique des énoncés qui forment un texte donné. Or, une telle opération est directement liée à la question de l'inférence : « l'automate "comprend" s'il est en mesure, à partir d'un énoncé quelconque, de produire, comme le ferait un locuteur humain, des énoncés qui sont vrais si l'énoncé de départ est vrai » [R. Martin 2004], i.e., de générer toutes les inférences valables pour cet énoncé. Nous poserons alors les postulats suivants :

- i. la structuration sémantique des textes peut être décrite au moyen de l'inférence
- ii. il existe au moins trois niveaux d'inférence : lexicale, énonciative et discursive

Les trois niveaux d'inférence correspondent à trois niveaux d'analyse : le niveau des phrases hors situation d'énonciation, le niveau des énoncés actualisés par un locuteur dans un contexte d'énonciation précis et le niveau du discours défini comme l'enchaînement cohérent d'une suite de phrases [Z. Harris 1952]. Nous monterons, dans cet article, que les trois niveaux d'inférence sont complémentaires et que les inférences énonciatives et discursives dépendent étroitement d'une description minutieuse des inférences lexicales. Ce premier niveau est la condition *sine qua non* pour que puissent être explicitées toutes les inférences du niveau supérieur.

Ce découpage en trois niveaux d'inférence rejoint la typologie de Michele Prandi [2004] qui établit une distinction binaire entre *inférence interne* qui participe à la construction du sens d'une prédication et *inférence externe* qui intervient, une fois que le sens lexical est construit, pour le souscrire et l'exploiter dans une situation de communication contingente. Les inférences lexicales seraient ainsi des inférences internes, liés à la notion de codage, en participant à déterminer l'interprétation des phrases, hors contexte. Par contre les inférences énonciatives et discursives sont déterminées par un contexte précis et déterminent les enchaînements dans une conversation ou un discours donné.

Après avoir présenté le corpus et le contexte applicatif, nous étudions les inférences lexicales, soit dans le cadre des prédicats élémentaires en les rapportant aux trois fonctions primaires qui définissent un emploi prédicatif donné (fonction prédicative, argumentale et actualisatrice), soit dans le cadre des prédicats de second ordre, à travers les phénomènes d'enrichissement inférentiel et de parataxe. Ensuite, nous montrerons comment les inférences énonciatives reposent sur des savoirs partagés entre locuteur et interlocuteur. Enfin, les inférences discursives, qui constituent le niveau supérieur de description des textes, nécessitent la prise en compte et la formalisation d'un ensemble de connaissances pragmatiques nécessaires à une interprétation globale et ciblée des textes.

1 Corpus d'étude et contexte applicatif

Le contexte applicatif de la description des inférences est la catégorisation sémantique automatique des réclamations clients. Les réclamations sont des documents textuels adressés aux entreprises, via plusieurs canaux d'interaction (e-mail, conversations d'agents virtuels, enquêtes de satisfaction, avis, forums, réseaux sociaux, etc.), dans lesquels les consommateurs formulent une demande explicite ou implicite qui appelle une réponse de la part de l'entreprise destinataire. Le système doit rapporter, dans un premier

temps, les documents à trois grandes catégories : (i) les questions, qui correspondent à des demandes majoritairement explicites ; (ii) les opinions, qui n'attendent pas nécessairement de réponse précise mais correspondent malgré tout à une demande, souvent implicite, de prise en considération, (iii) les injonctions, qui renvoient à une demande explicite ou implicite de réparation. Ces trois catégories correspondent à une analyse syntactico-sémantique qui peut avoir des configurations morpho-syntaxiques très variées. Par exemple, les phrases interrogatives peuvent désigner à la fois une question, une opinion ou une injonction.

Au plan formel, les réclamations sont des textes courts, très peu structurés, puisque les marques de ponctuation et de typographie sont peu respectées, avec une grande diversité thématique, et d'un niveau de langue assez familier. Ni les méthodes statistiques, ni les analyses qui procèdent par des ontologies ne peuvent apporter des résultats satisfaisants quant à l'extraction des principaux sujets qui constituent les motifs de contact¹, car la notion de fréquence ne caractérise pas ces textes, qui, par ailleurs, mélangent à la fois des vocabulaires techniques et un vocabulaire générique. Il y a donc avant tout une exigence méthodologique : les outils de traitement et le modèle théorique qui les conçoit doivent prendre en considération toutes les dimensions linguistiques, dans une approche intégrée où le sens ne se conçoit pas indépendamment du lexique, de la syntaxe et des paramètres énonciatifs.

Du point de vue du destinataire, les verbatim clients constituent une interaction, même implicite, nécessitant l'identification claire et rapide d'une action à mener, alors que le contenu du message présente souvent des caractéristiques qui vont à l'encontre de cet objectif :

- un message peut faire état de plusieurs sujets qui présentent chacun des demandes de type différent.
- les messages n'ont pas le même poids ; ce qui implique des les hiérarchiser ;
- les sujets peuvent être énoncés avec un niveau de précision et de complétude non homogènes, le discours le plus détaillé n'étant pas nécessairement relatif au sujet primaire.

La pertinence d'une réponse à un message dépend, en premier lieu, de la capacité à traiter le sujet, explicite ou implicite, qui constitue le fait générateur du message. Les méthodes utilisées pour identifier automatiquement les thèmes d'un message ne préjugent pas de la capacité à n'en retenir qu'un, celui sur lequel doit porter en priorité et a minima la réponse. Il convient de considérer deux enjeux : (i) identifier le sujet à traiter (le sujet primaire), et (ii) le type d'action à réaliser.

2 Inférences lexicales

Les inférences lexicales dépendent des structures prédicatives dans lesquelles trois catégories primaires d'unités lexicales sont décrites. L'analyse linguistique consiste alors à la fois à expliciter les inférences qui dépendent d'un emploi prédicatif donné et à décrire les mécanismes qui autorisent ces inférences, de sorte à pouvoir les traiter automatiquement.

2.1 Inférences et prédication élémentaire

Le cadre minimal d'analyse des inférences lexicales est la phrase élémentaire, définie comme une relation entre un prédicat de premier ordre et ses arguments. C'est cette relation qui permet d'interpréter les unités lexicales, et d'explicitement toutes les inférences qui leur sont rattachées. Cela signifie, d'une part, que les propriétés sémantiques sont subordonnées à la syntaxe, et, d'autre part, que les unités lexicales sont étudiées strictement du point de vue de leurs propriétés linguistiques. Le cadre de la phrase donne, par conséquent, un aspect formel à l'étude de l'inférence lexicale. À ce niveau, seule la structure prédicats-arguments est prise en compte, i.e., indépendamment de leur énonciation par un locuteur. Bien que ce niveau ne soit pas observable directement, car il s'agit d'une construction théorique, il constitue le niveau minimal d'analyse et d'information, qui sera emboîté, par la suite, dans des niveaux supérieurs (les énoncés et les discours). Si les problèmes d'ambiguïté et de figement ne sont pas résolus et les inférences lexicales non explicitées, l'automate ne peut remonter correctement aux inférences énonciatives et discursives.

La définition de la phrase qui fait appel aux deux procédures de linéarisation et d'actualisation permet de présenter une catégorisation des unités lexicales selon leur fonction dans la phrase. Les mots sont soit des prédicats, soit des arguments, soit des actualisateurs. Le prédicat est l'élément central qui sélectionne des arguments et des actualisateurs spécifiques. D'autres actualisateurs dépendent soit directement des arguments, soit de la relation prédicat-arguments. On distingue une signification lexicale pour le prédicat et les arguments, tandis que les actualisateurs ont simplement une signification grammaticale, et se chargent d'apporter les informations de temps, d'aspect, de voix, de personne, de modalité, etc. Selon cette typologie fonctionnelle des unités lexicales, on peut donc dire qu'il y a trois sortes d'inférences pour chaque fonction.

2.1.1 Inférences et fonction argumentale

Nous retiendrons ici le mécanisme de la métonymie qui joue un rôle central dans les inférences argumentales. Deux types de métonymies sont très présentes dans les réclamations clients et participent à leur structuration sémantique : la métonymie de type argumental à travers l'emploi de la première personne avec des prédicats non appropriés à des <humains> et la métonymie de type prédicatif qui donne lieu à des arguments hyperonymiques [T. Massoussi 2008b, 2009].

La métonymie argumentale explique plusieurs emplois de la première personne dans nos textes :

- (1) *J'ai été déconnecté*
- (2) *On était maintenu à quai pendant plusieurs minutes*
- (3) *J'ai été débité plusieurs fois*
- (4) *Je ne capte plus quand je suis chez-moi*

Ces emplois sont dérivés respectivement des emplois appropriés suivants :

- (1a) *Mon ordinateur a été déconnecté*
- (2a) *Notre train était maintenu à quai*
- (3a) *Mon compte bancaire a été débité plusieurs fois*
- (4a) *Mon téléphone portable ne capte plus.*

Une façon de rendre compte de ces métonymies consiste à postuler qu'elles résultent non pas d'une opération de réduction d'un schéma d'arguments plus long, mais d'un transfert de prédicat [Massoussi 2008a, 2008b]. Autrement dit, les verbes *déconnecter*, *être maintenu*, *débiter* et *capter*, qui sont appropriés respectivement à la classe des <machines>, des <moyens de transport ferroviaires>, des <comptes bancaires> et des <appareils téléphoniques>, sont transférés à la classe des <humains>, sur la base d'une corrélation syntactico-sémantique. Les prédicats qui se chargent de cette corrélation sont appelés prédicats relationnels ; ils sélectionnent des arguments strictement appropriés appartenant souvent au même domaine. De ce fait, ils codent très souvent une relation présupposée par les arguments respectifs et sont assimilables à des prédicats métalinguistiques.

- (1b) *Je (travaille, surfe sur internet) avec mon PC*
- (2b) *Nous avons monté à bord de ce train*
- (3b) *J'ai ouvert un compte bancaire auprès de cette banque*
- (4b) *J' (utilise, passe des appels) mon téléphone*

Si l'on reprend l'exemple (2), on peut dire que seuls les voyageurs qui se trouvent effectivement à bord du train peuvent énoncer qu'ils sont maintenus à quai. Par contre, si un groupe de personnes observe un train maintenu à quai, ils ne peuvent énoncer (2), du fait qu'ils ne deviennent pas partie prenante du moyen de transport et ne participent, pour ainsi dire, pas directement à la situation véhiculée par le prédicat [Kleiber 1994].

Pour traiter les inférences métonymiques dans un texte, nous construisons des dictionnaires électroniques et des grammaires locales qui représentent toutes les propriétés syntaxiques des différentes classes d'arguments sources, ici les <ordinateurs>, les <moyens de transport ferroviaires>, les <comptes bancaires> et les <appareils téléphoniques>, et nous isolons dans des sous-grammaires toutes les

propriétés qui caractérisent par métonymie les <humains>. Cette organisation de nos ressources présente l'avantage de pouvoir à la fois factoriser les propriétés partagées par les humains et les arguments qui leur sont corrélés, et distinguer les propriétés spécifiques aux arguments non humains. La création des sous-grammaires n'est pas aléatoire, puisque seules les propriétés syntaxiques qui vérifient la corrélation avec des humains sont prises en considération. Du point de vue de la génération automatique, on évite ainsi de produire des phrases malformées comme dans (7) à (9), et du point de vue de la reconnaissance, on évite d'avoir du bruit comme dans (10) où l'emploi du prédicat *en panne* avec la première personne n'est pas synonyme de son emploi avec les <appareils téléphoniques>.

- (5) *Mon téléphone ne capte pas* → *Je ne capte pas*
- (6) *Mon téléphone a été géo-localisé* → *J'ai été géo-localisé*
- (7) *Mon téléphone ne s'allume pas* → **Je ne m'allume pas*
- (8) *Mon téléphone s'éteint tout le temps* → **Je m'éteins tout le temps*
- (9) *Mon téléphone clignote sans arrêt* → **Je clignote sans arrêt*
- (10) *Mon téléphone est en panne depuis hier* → *#Je suis en panne depuis hier*

D'autres métonymies de la première personne nécessitent la prise en compte d'un contexte plus large, comme dans *Lors de mon dernier passage en caisse, je n'ai pas été reconnu*, qui implique que ma carte de fidélité n'a pas été reconnue. Indépendamment du niveau d'analyse, ces emplois sont à manipuler avec beaucoup de précaution, puisque l'enchaînement avec la première personne en surface est trompeur ; sémantiquement, cela renvoie à des entités différentes. De plus, les inférences métonymiques de ce type donnent des renseignements importants sur l'auteur du message (ici, l'analyse de l'inférence métonymique détecte que le locuteur possède une carte de fidélité, et que cette carte a probablement un problème de fonctionnement).

Un deuxième type de métonymie caractérise les messages des clients. Il s'agit des métonymies de type prédicatif qui, à partir d'un premier emploi prédicatif des unités lexicales (essentiellement des substantifs) génèrent un emploi argumental. Dans cet emploi argumental, le mécanisme de la métonymie donne lieu à des hyperonymes. Considérons les substantifs *commande* et *achat*. Leur polysémie est signalée par deux types d'emplois : un emploi prédicatif (de 11 à 12) et un emploi argumental (de 13 à 14) :

- (11) *J'ai passé une commande hier soir*
- (12) *J'ai fait beaucoup d'achats sur ce site*
- (13) *Cette commande est défectueuse*
- (14) *Je ne souhaite garder qu'un seul des deux achats*

Dans cet emploi argumental, *commande* et *achat* sont, à première vue, interchangeable avec les noms hyponymes de la même classe d'arguments, mettons la classe des <objets commerciaux>. Il n'y aurait donc pas de différence de sens entre les deux phrases (15) et (15a) : les deux présupposent une action initiale.

- (15) *Je n'ai toujours pas reçu ma commande*
→ *J'ai passé une commande*
- (15a) *Je n'ai toujours pas reçu mon (téléviseur, aspirateur, manteau)*
→ *J'ai commandé un téléviseur*

Pourtant, à y regarder de plus près, on s'aperçoit que seul (15) implique que le locuteur n'a pas reçu **la totalité** de sa commande, tandis que (15a) ne nous dit pas si le locuteur a commandé d'autres objets avec le téléviseur, et s'il a donc reçu la totalité de sa commande. Du point de vue des inférences, ces phrases ne sont donc pas équivalentes. Dans le contexte des messages clients, cela a un impact notamment pour les entreprises destinataires de ces messages, qui livrent en deux temps, et qui ont besoin de distinguer deux catégories de situations : les clients n'ayant pas reçu leur commande, et les clients n'ayant pas reçu un produit (seul ou dans un lot de commande). Cette situation pragmatique a un pendant linguistique puisque les arguments de la classe des <objets commerciaux> n'ont pas le même statut par rapport à l'inférence.

2.1.2 Inférences et fonction prédicative

Plusieurs inférences prédicatives sont de type « présuppositionnel », selon la typologie de Robert Martin [2004], qui définit la présupposition comme le cas où « une proposition p et son contraire ($non-p$) induisent que q est vrai [...]. Selon le « tiers exclu », p ne pouvant être que vrai ou faux, q est vrai dans tous les cas (puisqu'il est vrai si p est vrai et vrai également si p est faux) ». Ainsi, *Pierre veut divorcer* implique que *Pierre est marié* ; *Pierre ne veut pas divorcer* implique également que *Pierre est marié*. Dans les verbatim clients, la présupposition caractérise un grand nombre de prédicats, dont essentiellement ceux qui sélectionnent des prédicats argumentaux en position N1. C'est le cas par exemple des prédicats d'<annulation>, de <modification> ou de <rectification>. Ces prédicats présupposent une première occurrence du prédicat qui fait partie de leur schéma d'arguments. Cela explique l'emploi du possessif avec ces prédicats, comme dans les phrases ci-dessous :

- (16) *Je souhaite annuler ma commande*
- (17) *Je compte modifier mes achats*
- (18) *Je vais rectifier l'erreur qui figure sur mon compte client,*

qui présupposent respectivement :

- (16a) *J'ai passé une commande*
- (17a) *J'ai effectué des achats*
- (18a) *Il y a une erreur sur mon compte client*

Le test de la négation et de l'interrogation confirment qu'il s'agit de présupposés, puisque (16a), (17a) et (18a) sont toujours vraies, même si (16), (17) et (18) sont à la forme négative ou interrogative :

- (16b) *Est-ce que tu souhaites annuler ta commande ?*
- (17b) *Est-ce que tu comptes modifier tes achats ?*
- (18b) *Est-ce que tu vas rectifier l'erreur qui figure sur ton compte client ?*
- (16c) *Je ne souhaite pas annuler ma commande*
- (17c) *Je ne compte pas modifier mes achats*
- (18c) *Je ne vais pas rectifier l'erreur qui figure sur mon compte client*

En fonction de la classe du prédicat argumental, il peut y avoir des inférences « à plusieurs étages » comme dans (19) qui présuppose à la fois (19a), et (19b) :

- (19) *Je souhaite annuler le retour de ma commande*
- (19a) *J'ai passé une commande*
- (19b) *J'ai demandé de retourner cette commande*

On voit bien l'intérêt des inférences prédicatives, qui, comme les inférences argumentales, sont autant d'indicateurs permettant de positionner automatiquement le locuteur dans une phase du processus client (avant-vente, vente, après-vente). Le client qui souhaite annuler le retour de sa commande se situe nécessairement dans une étape d'après-vente, comme le démontrent les présupposés (19a) et (19b). Cette demande de retour implique que le client a été probablement en contact avec le service après-vente de l'entreprise. On voit alors comment s'articule l'inférence prédicative avec les inférences discursives. On ne peut supposer que le client fût en contact avec le service après-vente, si on ne détecte pas correctement l'inférence (19b).

2.1.3 Inférences et actualisation

Nous donnons ici quelques exemples d'inférences qui dépendent des actualisateurs employés avec les mêmes structures prédicatives. Nous retiendrons essentiellement les déterminants et les temps verbaux.

- **Déterminants :**

- **Articles définis / Indéfinis**

Chacun sait que l'opposition défini/indéfini repose entre autres sur le fait que seul le premier opère un présupposé d'existence. Cette opposition permet notamment de distinguer une injonction (et dans certains contextes une opinion négative) dans (20) et une opinion positive dans (21) :

(20) *Merci de mettre en rayon **de** bons produits*

(21) *Merci de mettre en rayon **vos** bons produits*

Dans (20), on ne présuppose pas l'existence des bons produits ; le locuteur ne fait que formuler une demande de mise en rayon de ces produits non réels qu'il juge bons. Par contre, (21) présuppose l'existence de ces bons produits, et le locuteur fait alors le constat d'une rupture de ces produits dans les rayons du magasin.

- **Déterminants quantifieurs**

La différence entre le défini et le quantifieur *certain*s dans les deux phrases ci-dessous est avant tout de nature inférentielle.

(22) *Le site est inaccessible*

(23) *Certains sites sont inaccessibles*

Contrairement au défini dans (22), l'emploi du quantifieur *certain*s dans (23) implique que d'autres sites sont accessibles, et qu'il s'agit probablement d'un problème de connexion avec le terminal utilisé (et non pas par exemple un problème du terminal lui-même ; ce que pourrait laisser croire la phrase (23), parmi plusieurs autres possibilités). Énoncés dans le contexte d'un message client destiné à un fournisseur d'accès à internet, ces deux phrases correspondent donc à deux situations différentes : (23) correspond à un problème de connexion, alors que (22) correspond à un problème du site web du fournisseur lui-même.

- **Opposition imparfait / présent de l'indicatif**

Si l'on compare les deux phrases (24) et (25), on observe que l'interprétation de l'imparfait est différente :

(24) *Votre site était clair et simple, et les rubriques étaient bien organisées*

(25) *Le séjour était agréable et le personnel de l'hôtel était accueillant*

Énoncée dans le cadre d'une enquête de satisfaction, la phrase (24) équivaut à une comparaison implicite entre deux versions du site ; l'imparfait infère alors que la version actuelle est à l'opposé de l'ancienne version. Par contre, dans (25), l'imparfait a uniquement une valeur temporelle : les prédicats appréciatifs *agréable* et *accueillant* sont simplement présentés comme une propriété d'entités entièrement situées dans le passé de l'énoncé. Dans (24), l'imparfait équivaut à une négation au présent de l'indicatif ; ce que l'on peut voir dans (26), contrairement à (25) où l'imparfait et le présent ont la même signification, si l'on met de côté la valeur temporelle et aspectuelle :

(26) *Votre site n'est plus clair et simple et les rubriques sont bien organisées*

(27) *Le séjour est agréable et le personnel de l'hôtel est accueillant*

2.2 Inférences et relations transphrastiques

Nous renvoyons aux travaux de Gaston Gross [2004, 2009] sur les relations transphrastiques pour :

- la remise en cause de l'analyse classique de la phrase complexe en termes de proposition principale et subordonnée;
- la critique de l'analyse des connecteurs comme des séquences figées inanalysables par des règles de combinatoire ;
- le rôle que jouent les substantifs et les adjectifs dans ces connecteurs ;
- et la hiérarchie prédicative entre prédication de second ordre et prédication élémentaire.

De la même façon que les restrictions de sélection qui existent entre un prédicat de premier ordre et ses arguments élémentaires, il existe des restrictions similaires entre les prédicats de second ordre et les prédicats élémentaires. Ces restrictions ont trait soit à la polarité (négative, positive) ou aux classes sémantiques. S'agissant de la polarité, on s'est, par exemple, aperçu, avec la projection des prédicats de *cause* sur un grand corpus (dix années du journal *Le Monde*) que le spectre nominal du verbe *causer* est constitué essentiellement d'événements négatifs (*mort* [602] ; *dégât* [223] ; *perte* [123] ; *dommage* [97] ; *préjudice* [88]), alors que *susciter* se combine préférentiellement avec des actions ou des états humains (*réaction* [864] ; *inquiétude* [551] ; *intérêt* [438] ; *interrogation* [396] ; *débat* [368]). Ces deux verbes sont présentés dans la tradition grammaticale comme des synonymes.

Si l'on considère les inférences dans les relations transphrastiques dans les messages clients, on observe que la relation causale est celle qui, de loin, est la plus structurante, à travers essentiellement deux mécanismes : l'enrichissement inférentiel et la parataxe.

2.2.1 Enrichissement inférentiel

Plusieurs connecteurs nécessitent une inférence pour avoir une interprétation causale. La procédure en question est un enrichissement inférentiel (König et Traugott 1988, Hopper et Traugott 1993). Dans les messages clients, la relation de cause peut donc être exprimée par plusieurs connecteurs initialement temporels (28) ou aspectuels (29 et 30) :

- (28) *Je souhaite annuler ma commande suite au retard que vous avez pris dans la livraison*
(29) *Chaque fois que je commande sur internet, j'ai des problèmes*
(30) *Toutes les fois que je voulais ajouter un article au panier, un message d'erreur apparaît*

2.2.2 Inférences et parataxe

La parataxe correspond à l'absence de connecteur qui relie deux phrases simples. Cette absence est palliée par l'inférence, uniquement quand il existe une appropriation entre les prédicats qui constituent les deux phrases simples. Il y a sans doute encore des recherches à mener sur l'appropriation, qui n'est pas toujours lexicale, mais peut être pragmatique. Si l'on considère les deux cas de parataxe ci-dessous, on remarque que l'existence d'une relation sémantique entre les deux phrases permet de restituer par inférence les relateurs de concession : *pourtant, bien que, etc.*

- (31) *Je n'arrive pas à me connecter... Le téléphone est (pourtant) parfaitement fonctionnel*
(32) *Je n'arrive pas à accéder à votre site ... La mise à jour s'est (pourtant) très bien déroulée*

Hors contexte, la phrase *Le téléphone est parfaitement fonctionnel*, correspond à une appréciation positive portant sur l'argument *téléphone*. L'inférence d'une relation de concession confère à cette phrase une valeur argumentative, puisqu'il s'agit alors pour le locuteur d'écarter l'éventualité que la cause du problème de connexion viendrait de son terminal. Dans les messages clients, ce phénomène est très récurrent, dans la mesure où le locuteur qui relate un dysfonctionnement dans un produit ou un service commercialisé par l'entreprise, tente d'obtenir une solution et veut souvent écarter la possibilité que le problème puisse être de sa responsabilité.

Avec la concession, la cause est la relation sémantique la plus présente et la plus structurante des interactions clients, et plus particulièrement les courriels adressés aux entreprises. Très souvent, les connecteurs de cause sont absents et c'est l'inférence qui prend le relais.

- (33) *Je n'arrive pas à me connecter sur mon espace client. J'ai oublié mon mot de passe.*
(34) *Je n'ai pas pu passer commande ... votre site était indisponible*

3 Inférences et énonciation

Si les inférences décrites jusque là dépendent exclusivement de l'articulation entre les prédicats et leurs arguments, celles qu'on examine ici font appel à des paramètres énonciatifs et interviennent, une fois que le sens lexical est construit, pour le souscrire et l'exploiter dans une situation de communication contingente. Plusieurs mécanismes sont en jeu comme la négation, la polyphonie, les savoirs partagés, etc. Le rôle de ces mécanismes et l'articulation entre les inférences lexicales et les inférences énonciatives sont des questions qui dépassent le cadre de ce texte. Nous nous intéressons ici seulement aux savoirs partagés, notion que nous empruntons à J-C. Anscombe dans ses travaux sur les stéréotypes et la généralité (cf. Bibliographie).

Dans notre corpus, plusieurs énoncés à la voix affirmative infèrent d'autres énoncés à la voix négative sur la base d'un savoir partagé par le locuteur et l'interlocuteur. C'est le cas des énoncés (33) et (34), qui infèrent respectivement (33a) et (34a) :

- (33) *J'ai perdu mon mot de passe*
(33a) *Je ne peux pas me connecter à mon espace client*

- (34) *J'ai trop peu de points de fidélité cumulés sur mon compte*
(34a) *Je ne peux pas renouveler mon téléphone portable*

Le passage de (33) à (33a) se fait moyennant le savoir partagé suivant : *Pour se connecter à son espace client, il faut un mot de passe.* De même, le passage de (34) à (34a) repose sur le savoir partagé : *Pour renouveler son téléphone portable, il faut beaucoup de points de fidélité.*

Dans les interactions clients, l'interprétation d'un dysfonctionnement ou d'un événement négatif dans le parcours du client s'explique par ces savoirs partagés. Cela repose très souvent sur une comparaison implicite entre une situation « normale » attendue et une situation présentée comme « perturbée » par le locuteur. Au plan morpho-syntaxique, plusieurs marqueurs prennent en charge l'expression des situations « perturbées » :

- des adverbes temporels de durée :

- (35) *Cela fait des heures que j'essaie de me connecter sur votre site*
(36) *Cela fait plusieurs minutes que j'essaie de joindre le service client*
(37) *Cela fait deux semaines que j'attends ma livraison*
(38) *Cela fait plus d'une heure que je tente de valider ma commande*

- des déterminants d'occurrences multiples :

- (39) *C'est ma (énième, troisième) relance à ce sujet*
(40) *Je vous ai déjà envoyé plusieurs mails*

- des adverbes aspectuo-temporels (*toujours, encore*) :

- (41) *Je n'ai toujours pas reçu ma commande*

- des semi-auxiliaires :

- (42) *Je suis contraint de vous appeler plusieurs fois pour suivre ma commande*
(43) *Je suis obligé de vous écrire plusieurs mails*

- des phrases interrogatives :

- (44) *Est-ce que j'aurai un jour ma commande*
(45) *Devrais-je attendre un déluge pour avoir ma commande*

Tous ces énoncés infèrent une situation « normale » attendue, mais non réalisée. Ainsi, *je n'ai toujours pas reçu ma commande* infère *Je devrais recevoir ma commande* ; *Cela fait plusieurs minutes que j'essaie de joindre le service client* infère *Je devrais joindre le service client sans attendre* ; *Je vous ai déjà envoyé plusieurs mails* infère *Je devrais avoir réponse dès le premier mail*, etc. Une grande partie des énoncés qui forment les messages envoyés par les clients d'une entreprise repose sur ces savoirs partagés, oppose les deux types de situation et véhicule, par conséquent, des réclamations voire des reproches implicites.

Parallèlement à ces savoirs partagés, certains énoncés reposent sur l'inférence d'un savoir acquis par une expérience directe. C'est par exemple le cas des prédicats d'appréciation employés avec le verbe *trouver* dans la tournure *Je trouve que*, et qui infèrent une expérience directe, laquelle peut être explicitée linguistiquement par des prédicats de perception visuelle, auditive, olfactive, etc. [J-C. Anscombe 2010 et à paraître].

- (46) *Je trouve que ce roman est excellent*
(47) *Je trouve que ce plat est délicieux*

L'interprétation des énoncés (46) et (47) n'est possible que si le locuteur a une expérience directe de l'objet jugé [Ducrot 1975]. L'énoncé (46) implique que le locuteur a lu le roman en question, et (47) implique que le locuteur a goûté du plat. Cette inférence explique les contrastes suivants

- (46a) **Je n'ai pas lu ce roman, mais je le trouve excellent.*
(46b) *J'ai lu ce roman, et je le trouve excellent.*
(47a) **Je n'ai pas goûté ce plat, mais je le trouve délicieux.*
(47b) *J'ai goûté ce plat, et je le trouve délicieux.*

Anscombe [2011] montre alors que « ce procédé nous fournit un excellent critère pour tester la présence d'un savoir acquis de façon sensorielle, comme on le voit sur :

Je vois la lune, et je la trouve très belle.
J'entends de la musique, et je la trouve très belle.
J'ai touché le tissu, et je le trouve très doux.
J'ai senti le parfum, et je le trouve très frais.
J'ai goûté le vin, et je le trouve très bon ».

La mise en place des tournures du type *je trouve que*, dans l'analyse des sentiments et des appréciations nous permet alors de déduire si le locuteur a vécu l'expérience de la chose jugée ; ce qui constitue donc un paramètre supplémentaire pour déterminer dans quelle phase du processus client se situe l'auteur du message.

4 Inférences discursives

Au niveau du discours, les inférences linguistiques interfèrent avec des connaissances extra-linguistiques. Le niveau d'analyse de ces inférences n'est plus la phrase ni l'énoncé, mais ce que l'on peut appeler des situations et qui se définissent comme la conjonction de paramètres linguistiques directement extraits du texte et de connaissances pragmatiques (ou connaissances « métier »). L'interprétation et la mise en place d'une situation font appel à des inférences linguistiques auxquelles s'ajoutent des connaissances extérieures au texte. Ces connaissances permettent notamment de moduler les premières inférences linguistiques et de leur ajouter de nouvelles significations.

Dans un message client, la diversité des sujets n'est pas strictement aléatoire, ni au plan thématique, ni au plan de l'occurrence. Le message client relève, en effet, d'un événement dans le processus de la relation client. L'analyse des situations au sein d'un message peut s'apprécier au regard de configurations types

liées à la structure même du processus de la relation client. Pour illustrer ce propos, considérons par exemple un message qui mentionne la notion de garantie d'un produit défectueux : il paraît vraisemblable (mais incertain) que l'émetteur du message ait déjà effectué son achat (comparativement à un message qui questionnerait sur la disponibilité d'un produit, par exemple). Dès lors, on peut faire l'hypothèse que le contexte du message est une situation d'après-vente pour laquelle nous savons qualifier le champ des configurations possibles relativement à la structure nominale du processus client, à cette étape du processus.

Par ailleurs, l'analyse inférentielle des situations nécessite une typologie linguistique des différents types de canaux d'interaction client. Il y a par exemple les récits d'expérience dans les avis, les conversations à deux actants avec les agents virtuels, les conversations à plusieurs actants dans des réseaux sociaux et des forums, les réponses à des questions dans les enquêtes de satisfaction, les discours directs dans les courriels, etc. Même si son texte fait état des mêmes prédicats, il est clair qu'un client ne se trouve pas dans la même situation selon qu'il répond à une enquête de satisfaction ou qu'il décide d'envoyer un courriel à l'entreprise.

5 Perspectives

Les perspectives de ce travail sont à la fois applicatives et théoriques. Les perspectives applicatives consistent à passer d'une catégorisation multiple des différents prédicats qui composent les messages clients à une catégorisation unique de l'ensemble des textes au moyen d'une représentation formelle des différentes inférences qui le structurent à tous les niveaux : phrastique, énonciatif et discursif. Le but est de construire des grammaires locales complètes pour le premier niveau d'inférence, et de spécifier les deux autres niveaux par des algorithmes de décision, en vue d'aboutir à une interprétation globale du message.

Sur le plan théorique, il s'agit de mettre en place un ensemble de critères linguistiques en vue d'une typologie des inférences qui montre tout le continuum qui existe en allant des inférences lexicales (implicatures conventionnelles chez Grice ou inférences internes chez Prandi) aux inférences discursives (implicatures conversationnelles chez Grice ou inférences externes chez Prandi). Plus largement, l'objectif est de mettre en place une véritable linguistique du discours qui s'appuie sur une description formelle du lexique en prenant en compte tous les paramètres énonciatifs.

Références bibliographiques

- Anscombe J-C. (2000). Un problème de sémantique lexicale: l'interprétation active/passive des adjectifs verbaux participes en position d'épithète. *Revue romane*, 45, 237-259.
- ANSCOMBRE J-C. (2006). Les objets de la polyphonie, *Le Français Moderne*, 74, 1.
- Anscombe J-C. (2010). Lexique et médiativité: les marqueurs pour le dire. *Cahiers de Lexicologie*, 1, 96, 5-33.
- Anscombe J-C. (2011). Généricité, analyticit  et propri t : une philosophie en langue ? *Cahiers de Lexicologie*, 99, 71-96.
- Ducrot, O. (1975). Je trouve que. *Semantikos*, 1, 62-88.
- Haillet P. P. (2002). *Le conditionnel en fran ais : une approche polyphonique*, Paris : Ophrys.
- Grice H. P. (1981). Presupposition and conversational implicature. P. Cole Radical Pragmatics ( ds.), New York, Academic Press, in H. P. Grice, *Studies in the way of words*, Harvard, Harvard University Press, 269-282.
- Gross G. (1995). Une s mantique nouvelle pour la traduction automatique: les classes d'objets. *La Tribune des Industries de la Langue et l'Information  lectronique*, 16-19.
- Gross G. (2009). *S mantique de la cause*, Peeters Louvain-Paris : Publications de la Soci t  de Linguistique de Paris.
- Gross M. (1995). Une grammaire locale de l'expression de sentiments. *Langue fran aise*, 105, 70-87.
- Harris Z. (1952). Discourse Analysis. *Language*, 28, traduit in *Langages*, 1969, 13, 8-45.

- Hooper, P. et Traugott, C. (1993). *Grammaticalization*. Cambridge University Press.
- Kleiber G. (1994). *Nominales. Essais de sémantique référentielle*, Paris : Armand Colin.
- König, E. et Traugott, E. C. (1988). Pragmatic strengthening and semantic change : the conventionalizing of conventional implicature. in W. Hüllen, R. Schulze (eds), *Understanding the Lexicon. Meaning, Sense and World Knowledge in Lexical Semantics*, Niemeyer, Tübingen, 110-124.
- Le Pesant D. et M. Mathieu-Colas (1998). Introduction aux classes d'objets. *Langages*, 131, 6–33.
- Martin R. (1976). *Inférence, antonymie et paraphrase*, Paris : C. Klincksieck.
- Martin, R. (2001). *Sémantique et automate*, Paris : PUF.
- Martin, R. (2002). *Comprendre la linguistique*, Paris : PUF.
- Martin R. (2004). Etiquetage sémantique du lexique français. *Le français moderne*, 72, 1, 14-26.
- Massoussi T. (2008). Représentation des inférences dans un dictionnaire électronique. *Cahiers de Lexicologie*, 93, 65-83.
- Massoussi T. (2009). Les prédicats partagés : entre catégorisation et inférence lexicale. *L'information grammaticale*, 122, 30-39.
- Pustejovsky J. (1995). *The Generative Lexicon*, Cambridge : The MIT Press.
- Prandi M. (2004). *The building blocks of meaning*, Amsterdam-Philadelphia: John Benjamins.
- Sperber D et Wilson D. (1989). *La Pertinence, Communication et cognition*, (traduit de l'anglais par Abel Gerschenfeld et Dan Sperber), Paris : les éditions du minuit.

ⁱ Les motifs de contact renvoient ici à l'ensemble des prédictions où le locuteur (qui est donc le client d'une entreprise destinataire du message) expose des problèmes rencontrés, formule des demandes ou affirme son opinion sur tel produit ou tel service. Ces prédicats constituent les motifs pour lesquels il a décidé d'écrire son message.