

Influence de la situation lors de la résolution des anaphores dans le dialogue

Frédéric Landragin

Thales Research & Technology, Palaiseau
frederic.landragin@thalesgroup.com

Résumé

La résolution des anaphores dans les systèmes de dialogue homme-machine s'inspire généralement des modèles et des algorithmes développés pour le texte. Or le dialogue met en jeu une situation, c'est-à-dire un environnement physique immédiat et des événements dont la perception est partagée par les interlocuteurs. Cette situation peut servir d'ancrage à des expressions référentielles dites « anaphores à antécédents non linguistiques ». L'attribution de référents à de telles expressions s'avère difficile pour deux raisons : premièrement les facteurs situationnels sont nombreux et peu explicites ; deuxièmement des ambiguïtés peuvent apparaître entre de possibles antécédents situationnels et de possibles antécédents linguistiques. Nous proposons ici un modèle clarifiant l'intervention des facteurs situationnels et permettant leur prise en compte lors de la compréhension des expressions référentielles potentiellement anaphoriques. En intégrant la notion de saillance valable à la fois pour les aspects situationnels et linguistiques, nous montrons comment utiliser des scores numériques pour gérer les interférences entre hypothèses situationnelles et linguistiques.

Mots-clés : co-référence, anaphore, contexte, perception visuelle, saillance.

Abstract

Anaphora resolution in human-machine dialogue systems is often based on models and algorithms that were designed for text. But dialogue happens in a situation, *i.e.*, an immediate physical environment and events that are perceived simultaneously by the participants. This situation can be at the origin of referential expressions called "anaphora without linguistic antecedents". Attributing some referents to such expressions is difficult for two reasons. First, because situational factors are numerous and implicit. Second, because ambiguities can appear between a potential situational antecedent and a potential linguistic antecedent. In this paper we propose a model that clarifies how situational factors work and how they can be taken into account when interpreting anaphoric expressions. By integrating the notion of saliency that is common to both situational and linguistic aspects, we show with numeric scores how the interference between the situation and the language can be solved.

Keywords: coreference, anaphora, context, visual perception, saliency.

1. Introduction

Bien qu'il existe d'innombrables travaux en linguistique et en traitement automatique des langues sur la résolution des anaphores, la plupart se focalisent sur l'anaphore dans le texte, et très peu cherchent à replacer l'anaphore dans des situations de dialogue où le contexte prend des formes variées et complexes. Deux exemples de telles situations de dialogue sont d'une part la communication humaine de face à face, où les interlocuteurs partagent un espace visuel et peuvent conforter leurs énoncés linguistiques par des gestes conversationnels, et d'autre part le dialogue homme-machine reposant sur une scène visuelle affichée à l'écran.

Dans les deux cas, la résolution d'une anaphore ne repose pas sur le seul matériau linguistique, mais aussi sur la nature du contexte visuel et sur la survenue d'événements dont

la perception est partagée par les interlocuteurs. Ces événements dont l'origine n'est pas linguistique peuvent mettre en jeu l'un des interlocuteurs (c'est le cas d'un geste produit par l'un d'eux) ou être totalement extérieurs au dialogue, pouvant alors diriger celui-ci dans une nouvelle direction. Dans la littérature, on trouve quelques rares exemples détaillés de telles situations. L'un des premiers est peut-être celui de (Isard, 1975) qui met en scène un enfant tendant la main à travers les barreaux d'une cage où est enfermé un lion, et à qui on dit : « *attention, il risque de te mordre !* ». Dans cet énoncé, le pronom « *il* » constitue un cas typique d'anaphore à antécédent non linguistique. L'antécédent est ici le lion, et son identification ne peut se faire qu'à travers l'action de l'enfant, action perçue par le locuteur et exploitée dans son énoncé. Le fait que l'action ne provient pas du locuteur justifie ici l'emploi du terme anaphore, par opposition à celui de déixis qui regroupe les références déictiques (où le locuteur s'aide intentionnellement d'un geste pour désigner un référent, par exemple « *il* » associé à un geste indiquant le lion) et les indexicaux (« *je* », « *ici* », « *maintenant* »). À partir de l'exemple d'Isard, il est très facile d'imaginer d'autres situations, par exemple la phrase « *il joue vraiment bien* » énoncée lors d'une représentation théâtrale, à un moment où l'acteur principal est particulièrement éclairé sur la scène, ou encore la phrase « *il manque de souplesse* » énoncée par une personne en train de se caler dans un fauteuil. Pour ce dernier cas, il est même possible d'envisager des ambiguïtés apparentes entre antécédent situationnel et antécédent linguistique, par exemple si la conversation courante porte sur le caractère de quelqu'un : soit l'événement situationnel est prépondérant et « *il* » reprend l'antécédent fauteuil, soit le topique du dialogue est prépondérant et « *il* » reprend l'antécédent humain. L'exemple d'Isard peut aussi être transformé de manière à y ajouter une ambiguïté potentielle, si on considère par exemple que le locuteur et l'enfant viennent de parler du chien qui trotte à leur côté, avec un énoncé tel que « *Médor a l'air d'avoir faim* ».

Dans cet article nous voulons faire le point sur les facteurs linguistiques et situationnels qui interviennent lors de la résolution des anaphores, de manière à clarifier les rôles d'un module de résolution de la référence dans un système de dialogue homme-machine. Même si les exemples d'ambiguïtés que nous avons donnés peuvent sembler marginaux, nous pensons au contraire qu'ils illustrent des situations de communication fréquentes dans la vie courante, et qu'un système capable de les analyser correctement serait bien plus compréhensif et efficace qu'un système limité aux seuls aspects linguistiques. Ces exemples nous serviront donc de fil directeur dans notre étude des facteurs linguistiques et situationnels, ainsi que dans la confrontation de ces deux types de facteurs. Il ne s'agit pas de proposer un nouveau modèle de résolution des anaphores, mais de fournir des arguments pour appréhender ce problème d'une manière plus générale et plus appropriée au dialogue. Du fait de notre focalisation sur les facteurs permettant d'identifier un antécédent, nous n'explorerons pas plus les formes d'anaphores et nous resterons dans le cadre de l'anaphore pronominale. Après une section portant sur les facteurs linguistiques utilisés dans les modèles existants et envisagés dans notre modèle, puis une section dédiée à notre classification des facteurs situationnels, nous présenterons notre première proposition de modèle pour la gestion des interférences entre langage et situation lors de la résolution des anaphores. Cette proposition restera théorique dans la mesure où il n'existe pour le moment aucun moyen de procéder à une évaluation sur corpus. Les facteurs situationnels n'ont en effet pas de trace linguistique, et ne sont donc inclus de manière objective dans aucun corpus, du moins sous une forme qui soit propice à un traitement automatisé. Il faudrait en effet disposer à chaque fois d'une description exploitable de la scène visuelle et des événements qui y ont lieu, un peu à la façon des didascalies dans les pièces de théâtre (qui restent pour leur part trop subjectives et donc inexploitables). Afin de montrer son intérêt, nous montrerons comment notre proposition se matérialise avec les exemples donnés ci-dessus, puis nous donnerons quelques perspectives pour l'améliorer.

2. Facteurs linguistiques pour la résolution des anaphores

La résolution des anaphores, et en particulier des anaphores pronominales, a fait l'objet de nombreux travaux en linguistique et en traitement automatique des langues. Les méthodes varient entre la prise en compte de contraintes surtout syntaxiques (Grosz *et al.*, 1995 ; Kamp et Reyle, 1993), la prise en compte de contraintes couvrant la syntaxe et la sémantique de la phrase (Hajičová *et al.*, 1995), la mise en place de scores numériques traduisant les saillances relatives des antécédents potentiels (Alshawi, 1987 ; Lappin et Leass, 1994), ou encore la mise en place de procédures calculatoires variées (Mitkov, 1998), plus proches d'un calcul informatique facilement réalisable mais plus éloignées des fondements linguistiques. Dans l'ensemble, il s'agit de partir du matériau textuel et d'en tirer des indices pour bloquer certains antécédents potentiels. À partir de là, soit il ne reste plus qu'un antécédent potentiel et le problème est résolu, soit il reste plusieurs candidats. Les indices syntaxiques et sémantiques peuvent alors être exploités pour classer ces candidats et retenir le meilleur. Dans les approches les plus orientées sur la cognition humaine, des hypothèses sont faites sur la mémoire du lecteur, avec des aspects tels que l'accessibilité et la récence de mention. Ces hypothèses ont également pour but de classer les candidats et de retenir l'antécédent le plus accessible, le plus récent, le plus saillant d'un point de vue cognitif. Elles cherchent à ne pas trop s'éloigner du matériau linguistique sur lequel elles se fondent, mais elles ouvrent néanmoins les portes d'un domaine très vaste qui pourrait être d'un grand apport : il apparaît de plus en plus possible d'exploiter non seulement les aspects mémoriels et attentionnels, mais aussi des aspects cognitifs plus complexes tels que la représentation mentale des entités du discours et les raisonnements sur ces représentations.

Dans le but de poser des pistes possibles pour notre problème d'ambiguïté entre antécédents situationnel et linguistique plus que dans le but de proposer un véritable état de l'art sur la résolution des anaphores, nous pouvons synthétiser de la manière suivante :

1. Identification et traitement des contraintes linguistiques, en particulier celles relatives au genre et au nombre, avec éventuellement quelques contraintes sémantiques. Dans l'exemple de l'énoncé « *attention, il risque de te mordre !* » prononcé peu de temps après « *Médor a l'air d'avoir faim* » et, pourquoi pas, « *tu as vu le gardien du zoo ?* », le pronom personnel est au masculin et au singulier, contraintes que les antécédents potentiels « *Médor* » et « *le gardien du zoo* » vérifient bien. D'autre part, « *il* » étant susceptible de « *mordre* », une contrainte supplémentaire porte sur la nature animale ou humaine de l'antécédent, contrainte que chien et gardien vérifient également. Nous noterons que cette dernière contrainte peut être relâchée, d'une part parce que dans des contextes précis, un objet inanimé est susceptible de mordre (on pensera par exemple à une machine industrielle ou à des pinces crocodiles), et d'autre part parce que l'usage de métaphores peut transgresser ce type de contraintes.
2. Comme deux antécédents restent possibles, il s'agit maintenant de les classer à partir des indices linguistiques, en particulier la récence de mention, la fonction grammaticale, le rôle thématique, le topique, etc. (Sidner, 1979). Sans entrer dans les détails des facteurs privilégiant tel ou tel antécédent (pour une revue, cf. Landragin, 2004), nous noterons que « *Médor* » est privilégié par sa récence de mention (en supposant que l'énoncé le concernant ait été prononcé juste avant l'anaphore pronominale), par sa fonction grammaticale sujet, par sa position en tête de phrase, etc. Il est vraisemblable que, d'une manière générale dans les systèmes existants, ce soit effectivement « *Médor* » qui soit choisi comme antécédent privilégié.

3. Une autre façon de considérer le problème est que le classement précédent ne tient pas compte des aspects situationnels et par conséquent n'est pas valable. La situation comportant une action de l'enfant vers le lion enfermé dans sa cage, les participants de cette action sont à prendre en compte au même titre que les entités de discours qui ont été mentionnées. Or le lion vérifie lui aussi les contraintes linguistiques de genre, de nombre et d'être animé susceptible de mordre. De plus, « *attention* » constitue un indice linguistique important dans la mesure où il est souvent utilisé pour porter l'attention de l'interlocuteur vers un nouvel élément du contexte ou vers un changement de contexte (pour une étude de cette notion de changement de contexte et des indices associés, cf. Isard, 1975 ; Beun et Cremers, 1998).

3. Facteurs situationnels pour la résolution des anaphores

Parmi les diverses applications du traitement automatique des langues, c'est essentiellement dans le dialogue homme-machine que l'on peut trouver une prise en compte des caractéristiques de la situation pour la résolution des références et plus précisément pour la résolution des anaphores. D'une manière générale, plusieurs catégories peuvent être distinguées selon que le système est capable ou non de gérer des événements extérieurs à l'interaction en plus des événements utilisateurs, qu'il est capable ou non d'ordonner l'importance de ces événements, etc. Typiquement, dans le dialogue homme-machine spontané : les événements utilisateurs sont privilégiés aux autres événements car ils correspondent directement à la finalité du système (satisfaire l'utilisateur) ; le contexte visuel courant est privilégié à l'historique du dialogue car il est a priori plus lié à la tâche en cours ; la saillance et les facteurs permettant de classer des antécédents potentiels ne sont exploités qu'en dernier recours car ils restent hypothétiques. Au contraire, dans le dialogue opérationnel, c'est-à-dire pour lequel l'utilisateur est un opérateur qui connaît et maîtrise ses rôles comme c'est le cas dans la majorité des systèmes Thales (systèmes d'aide à la décision, interfaces pour le contrôle aérien ou la surveillance maritime, etc.), les événements ou actions de l'utilisateur et les événements externes sont traités avec le même niveau d'importance, chacun d'entre eux activant une procédure bien spécifiée, mais avec souvent une indépendance des processus qui ne permet pas d'autoriser la moindre coréférence ou anaphore. Dans ce type de système, les hypothèses sur les facteurs linguistiques ou cognitifs sont généralement proscrites, le système préférant des mentions très explicites des référents, quitte à demander des reformulations pour éviter toute interprétation aléatoire (et donc potentiellement dangereuse).

Toujours dans le but d'identifier des pistes pour l'interprétation de nos exemples ambigus, nous proposons la classification suivante des types d'événements à prendre en compte :

1. Événements liés à l'énonciation. Ces événements concernent le locuteur et incluent l'ensemble des gestes conversationnels produits conjointement à la parole, ainsi que les postures et les attitudes. On trouve ainsi d'une part les gestes communicatifs qui participent au dialogue et aux contenus sémantiques échangés, avec comme exemples les gestes co-verbaux (déictiques, illustratifs, expressifs, paraverbaux), les quasi-linguistiques qui peuvent se substituer à la parole pour exprimer un message (signes conventionnels constituant un lexique d'une centaine de gestes par culture), et les gestes synchronisateurs tels que les gestes de passage ou de maintenance de tour de parole. Chacun de ces gestes constitue un événement lié à l'énonciation susceptible d'être repris ensuite par une anaphore, ne serait-ce qu'à un niveau métalinguistique (« *vous dites ?* »). On trouve d'autre part les gestes extra-communicatifs qui ne participent pas vraiment au dialogue mais jouent plutôt le rôle d'indicateurs sur les états intentionnels ou les émotions du

locuteur. Il s'agit des gestes automatiques, centrés par exemple sur le corps ou sur des objets (comme la manipulation d'un stylo, geste automatique très fréquent), pouvant ainsi introduire l'objet comme nouveau référent dans le dialogue, sans antécédent linguistique (« *tu vas le casser* »). Il s'agit également des gestes praxiques et ludiques tels que les gestes utilitaires, par exemple tous les grattements. Là encore, la reprise anaphorique est possible, même si elle détourne le dialogue de son sujet courant. Concernant les postures et les attitudes, nous citerons le rapprochement ou l'éloignement du locuteur par rapport à son interlocuteur, les positions du corps qui expriment un certain relâchement ou une certaine impatience, ou encore les impressions transmises par le regard. Tout ceci est susceptible de rendre saillant un élément et de rendre ainsi possible une anaphore.

2. Événements liés à une tâche effectuée conjointement au dialogue. Ces événements concernent à la fois le locuteur et l'interlocuteur, et incluent le fait qu'ils marchent côte à côte tout en parlant, ou qu'ils se trouvent dans une salle de réunion avec un tableau blanc leur servant de support, etc. Avec l'exemple d'une réunion de travail, la tâche effectuée conjointement au dialogue peut être la rédaction coopérative d'un texte (les échanges coopératifs étant à l'origine du dialogue), le dessin d'un schéma technique au tableau, ou encore la préparation en temps réel d'un compte rendu de réunion. Là aussi, des phrases telles que « *ça avance ?* » ou « *il a l'air incomplet, on devrait ajouter un paragraphe* » peuvent être énoncées sans aucun antécédent linguistique et constituer des exemples d'anaphores à antécédents situationnels (avec « *il* » comme avec « *ça* », le problème est le même). On peut également inclure dans cette catégorie de facteurs les aspects sociologiques sous-jacents à un dialogue entre par exemple un employé et son chef, la tâche effectuée conjointement au dialogue consistant à ne pas perdre la face, à donner bonne impression, à paraître sociable, etc.
3. Événements liés à une tâche que le dialogue empêche ou a mis en attente, mais qui restent prévisibles pendant le dialogue. C'est le cas typique d'un dialogue engagé lors de l'attente d'un train ou juste avant le début d'une réunion de travail, un tel dialogue pouvant d'ailleurs se matérialiser par des échanges relativement courts et indépendants, par opposition à un dialogue composé de nombreux enchâssements ou sous-dialogues. Ces événements concernent les deux interlocuteurs et leur survenue potentielle reste active dans leur mémoire. L'anaphore est très facile : même si le train n'a pas été mentionné, des énoncés tels que « *ça y est* », « *le voilà* » ou « *il n'est en retard que de 5 minutes* » s'interprètent aisément. Contrairement à la catégorie suivante, le train ne constitue pas un nouvel élément dans la situation de communication dans la mesure où il est bien présent dans l'esprit des interlocuteurs.
4. Événements externes au dialogue et totalement imprévisibles dans le contexte de communication. Ces événements concernent également les deux interlocuteurs, et sont à l'origine d'une nouvelle situation de dialogue dans la mesure où ils peuvent rompre la discussion en cours et l'orienter dans une direction différente. C'est l'exemple du chef qui apparaît en pleine pause café, déclenchant ainsi la reprise du travail, ou encore celui de l'alerte incendie. La saillance du nouvel élément permet ici aussi l'utilisation d'un pronom anaphorique.

Les événements ont été classés selon leur lien avec le dialogue en cours, mais en aucun cas selon leur importance pour leur prise en compte dans le processus d'attribution d'un référent à un pronom. Il est en effet difficile d'attribuer des niveaux d'importance ou de priorité entre des aspects si disparates. L'alerte incendie correspond généralement à un niveau d'urgence qui devrait favoriser sa saillance dans tout énoncé prononcé suite à son déclenchement, ce cas

montrant que le degré d'importance dépend non pas de la catégorie de l'événement mais du contenu sémantique qu'il véhicule.

Dans l'information véhiculée par l'événement se trouvent les objets et participants de l'événement. Ce sont ces entités plus que l'événement lui-même qui constituent des antécédents situationnels potentiels pour le pronom anaphorique (l'événement pouvant ceci dit être un antécédent potentiel privilégié au terme « *ça* » présent dans certains de nos exemples). C'est ainsi que le train constitue un antécédent potentiel. C'est aussi de cette façon que l'on peut préparer l'interprétation de l'exemple d'Isard « *attention, il risque de te mordre* » : l'événement à la base de l'énonciation est le mouvement de l'enfant vers la cage où est enfermé le lion. Le principal participant est ainsi l'enfant, mais la cage, le lion, ou encore les barreaux de la cage et les mâchoires du lion sont des entités concernées par l'événement et constituent des référents potentiels. La détermination des antécédents possibles passe ainsi par la modélisation du micro-monde autour de l'événement. Du fait des contraintes syntaxiques et sémantiques de l'énoncé, seules sont activées les entités au masculin singulier susceptibles de mordre l'enfant, ce qui permet de retenir le lion (et exclut l'enfant lui-même).

4. Interférences entre les aspects linguistiques et situationnels

4.1. Paramètres permettant de faire un choix

Nous nous intéressons maintenant aux cas d'ambiguïté entre antécédent linguistique et antécédent situationnel. Nous disposons pour le moment de critères permettant de déterminer les antécédents linguistiques potentiels, nous avons de même des critères pour déterminer les antécédents situationnels potentiels, mais rien pour favoriser l'une ou l'autre façon d'interpréter. Ce problème a déjà été soulevé, en particulier pour la confrontation de la saillance linguistique et de la saillance visuelle, mais sans solution immédiate (Landragin, 2004). D'une manière générale, quelques résultats fondamentaux en psychologie cognitive montrent que l'on tend à favoriser les aspects visuels par rapport aux aspects linguistiques et mémoriels : les événements perceptibles visuellement sont plus percutants que ceux perceptibles auditivement, les images sont mieux mémorisées que les sons (Weil-Barais, 1993). Nous noterons au passage que les antécédents linguistiques, du fait de leur énonciation préalable au cours du dialogue, appartiennent à l'historique du dialogue et relèvent donc plutôt des aspects mémoriels que linguistiques. Mais cela ne nous permet pas de poser définitivement que les antécédents situationnels seront toujours à privilégier par rapport aux antécédents linguistiques. Au contraire, l'ordre d'importance est à déterminer au cas par cas.

Dans ce but, nous proposons dans la liste suivante un ensemble de paramètres permettant de choisir un ordre de préférence :

1. La séquence temporelle des événements et des mentions linguistiques. L'événement situationnel ou énonciatif le plus récent est favorisé dans la mesure où il est probablement plus présent dans l'esprit des interlocuteurs. Récence ne veut cependant pas dire saillance (il faut entrer dans des considérations plus sémantiques pour cela), et ce critère ne constitue qu'un paramètre parmi d'autres qu'il s'agira de mettre ensemble.
2. Les aspects prosodiques lors de la production du pronom anaphorique. Un premier aspect est le rythme : toute rupture dans un rythme régulier, toute pause est considérée comme significative et doit être analysée en tant que telle. Si la mention de l'anaphore suit directement la mention de l'antécédent linguistique, sans rupture de continuité, et qu'un

événement situationnel se produit en même temps que la mention de l'antécédent linguistique, rien ne peut être décidé et l'ambiguïté est maintenue. Si par contre, dans la même configuration, une pause précède la mention de l'anaphore, la rupture de continuité entre antécédent linguistique et anaphore conduit à privilégier l'antécédent situationnel. Un deuxième aspect regroupe l'intonation, les accents : si les deux énoncés diffèrent fortement dans leur prosodie, alors cette rupture de continuité prosodique conduit à privilégier l'antécédent situationnel. C'est le cas lorsque la phrase « *Médor a l'air d'avoir faim* » est énoncée calmement alors que « *Attention, il risque de te mordre !* » traduit une certaine inquiétude avec une hauteur plus importante ou un accent marqué sur « *attention* ». Au contraire, si les deux phrases sont énoncées avec la même intonation, on privilégiera l'antécédent linguistique par rapport à l'antécédent situationnel dont la prise de connaissance n'est pas révélée prosodiquement.

3. Les aspects lexicaux, avec le fameux « *attention* » dont le rôle est multiple. Présupposant que l'attention de l'interlocuteur devrait se focaliser sur quelque chose de plus important que la tâche courante, ce terme peut tout d'abord traduire la survenue d'un événement exceptionnel, et favoriser de ce fait les aspects situationnels. Avec une prosodie adéquate, il réfère de manière privilégiée à la situation immédiate et non à l'historique du dialogue qui n'est affecté par aucune espèce d'urgence. De plus, sa nature lexicale tend à faire l'hypothèse d'une rupture dans le dialogue, c'est-à-dire à privilégier les événements les moins liés au dialogue en cours, et par conséquent les aspects situationnels, surtout si ceux-ci appartiennent aux catégories 3 (événements liés à une tâche que le dialogue a mis en attente) ou 4 (événements externes au dialogue).
4. Les aspects syntaxiques. La syntaxe peut également apporter des indices de rupture de continuité permettant d'identifier un écart entre la mention de l'antécédent linguistique et celle de l'anaphore, et privilégiant de ce fait l'antécédent situationnel. Un manque de cohésion textuelle peut ainsi être un indice déterminant. Un autre indice jouant dans le même sens est la présence d'un énoncé sans prédication « *oh la belle bleue !* », d'une expression référentielle isolée, ou encore de toute expression vocative qui ramène préférentiellement à la situation.
5. Les aspects sémantiques. Au niveau de la signification des énoncés, on peut tout d'abord relever l'ensemble des expressions ramenant explicitement à la situation. Si ce n'était pas tout à fait le cas avec « *attention* », c'est en revanche clairement le cas avec « *regarde, voici le gardien du zoo* » ou avec « *tu as vu ? il t'a mordu* ». Parallèlement aux aspects syntaxiques, on peut ensuite mentionner les ruptures de cohérence qui peuvent indiquer de manière significative la survenue d'un événement extra-dialogique. À titre d'illustration, en considérant que la succession classique de deux phrases est celle où le rhème de la première est repris comme thème de la seconde (Caron, 1989), tout manquement à cette convention peut être un indice significatif de la prise de conscience d'un événement extérieur au dialogue. Si par exemple l'expression anaphorique est incluse dans un deuxième rhème, c'est qu'elle reprend sans doute un antécédent situationnel plutôt que l'antécédent rhème de la première phrase.
6. Les aspects pragmatiques. La rupture de continuité sur laquelle nous avons fondé notre approche peut s'appliquer aussi aux forces illocutoires (en les réduisant aux trois catégories très différentes que sont l'assertion, la question et l'ordre, de manière à mieux rendre compte des variations) : si l'énoncé contenant l'anaphore diffère des précédents par une force illocutoire inadéquate ou imprévisible, avec par exemple un acte de type ordre dans une série de questions ou une série d'assertions (narration d'une histoire), alors il peut être révélateur de la prise en compte d'un événement extérieur. En reprenant notre

exemple, c'est le cas de la suite d'énoncés « *Médor a l'air d'avoir faim. Fais attention, il risque de te mordre* » par rapport à « *Médor a l'air d'avoir faim. Il risque de te mordre* », l'anaphore textuelle étant légèrement plus manifeste dans ce dernier cas. Les aspects prosodiques qui leur sont indissociables devraient cependant suffire à rendre compte de ces phénomènes.

7. Les aspects cognitifs et les hypothèses sur les états mentaux des interlocuteurs. Pour terminer cette liste, il est intéressant d'évoquer les états intentionnels, du moins lorsque le dialogue homme-machine est suffisamment finalisé pour rendre pertinente l'émission par le système d'hypothèses sur ces aspects. Ainsi, lorsque la tâche est prédictible et qu'il est possible d'identifier un événement attendu (tel que l'arrivée du train), le système peut mettre en avant l'hypothèse de référent associée.

4.2. Exploitation des paramètres

Nous avons donc une liste d'indices permettant de privilégier l'interprétation linguistique ou l'interprétation situationnelle dans le cas d'une ambiguïté. Combinée avec les contraintes syntaxiques et sémantiques liées à l'analyse de l'énoncé contenant l'anaphore, cette liste devrait permettre d'aboutir à un classement des antécédents potentiels. Or qui dit classement dit quantification, du moins dans la perspective algorithmique qu'est celle du dialogue homme-machine. Comme nous l'avons décrit de manière détaillée dans (Landragin, 2005) au sujet de la quantification de la saillance, l'utilisation de scores numériques permet des comparaisons immédiates entre diverses hypothèses et constitue une stratégie efficace en sémantique computationnelle. La procédure à base du « principe de primordialité » (favoriser un candidat sur la base de la satisfaction d'un paramètre important), et du « principe de singularité » (favoriser un candidat sur la base d'une singularité dans la satisfaction de paramètres, c'est-à-dire favoriser le candidat qui est le seul à satisfaire le plus grand nombre de paramètres) s'applique clairement à notre problème de résolution des anaphores en situation de dialogue, les scores étant calculés par le ratio : nombre de paramètres satisfaits sur nombre total des paramètres intervenant dans la situation en question.

Ainsi, si nous reprenons encore une fois l'exemple basé sur la suite d'énoncés « *Médor a l'air d'avoir faim – Attention, il risque de te mordre !* » pour lequel nous avons analysé quelques variantes prosodiques et syntaxiques, nous pouvons faire le récapitulatif suivant : au niveau de la séquence temporelle des événements, rien ne peut être conclu ; au niveau des aspects prosodiques, nous garderons le cas de figure où la prosodie penchait pour l'antécédent situationnel, ce qui fait un premier point pour celui-ci ; au niveau lexical, « *attention* » favorise également l'antécédent situationnel (1 point supplémentaire) ; au niveau syntaxique, les deux énoncés peuvent être considérés comme présentant une certaine cohésion, ce qui attribue un point à l'antécédent linguistique ; au niveau sémantique, une certaine cohérence est également constatée et penche pour l'antécédent linguistique ; au niveau pragmatique, nous garderons l'apport directif de « *attention* » que notre choix de prosodie confirme et nous considérerons que l'antécédent situationnel est favorisé ; enfin, au niveau cognitif, nous préférons ne pas faire intervenir d'hypothèses sur les intentions du locuteur et nous laissons les scores dans leur état, c'est-à-dire 3 pour l'antécédent situationnel (le lion) et 2 pour l'antécédent linguistique (le chien). Du fait de la présence d'un événement situationnel et du score maximum qui lui est attribué, le système décide d'attribuer le lion comme référent au pronom anaphorique, ce qui bien entendu nous convient très bien.

5. Conclusion et perspectives

Comme toujours, le dialogue complique les méthodes et algorithmes prévus pour le texte, dans la mesure où il apporte tout un ensemble de paramètres supplémentaires, et, qui plus est, de paramètres hétérogènes couvrant plusieurs modalités de communication et plusieurs niveaux de communication, du plus explicite au plus implicite. Appliquer directement les travaux élaborés pour le texte ne nous semble ainsi pas pertinent, et il nous apparaît fondamental de clarifier au préalable la nature des phénomènes apparaissant en situation de dialogue. C'est ce que nous avons fait ici en nous focalisant sur un problème délicat, celui des interférences entre le matériau linguistique et le matériau situationnel pour la résolution des anaphores. Nous espérons avoir contribué à clarifier ce terrain peu exploré en proposant des classifications systématiques des facteurs situationnels, des facteurs linguistiques (dans une moindre mesure car ce sujet a été largement étudié par ailleurs), et surtout des paramètres permettant d'ordonner les importances relatives de ces facteurs hétérogènes.

Si nous passons directement à la conclusion, sans la classique validation à base de corpus, c'est qu'une telle évaluation est pour le moment quasiment impossible. Il n'existe en effet aucun corpus intégrant une description exploitable des événements physiques et du contexte visuel dans des situations de dialogue qui soient suffisamment riches pour intégrer en nombre suffisant les phénomènes qui nous intéressent. Comme nous le montrent nos discussions de tous les jours, ces phénomènes sont pourtant bien présents dans la vie courante et c'est pourquoi il nous a semblé important de les décortiquer. Reste à constituer et à analyser un corpus adéquat, en partant par exemple des textes et didascalies de théâtre, ce qui constitue la principale perspective de ce travail.

Références

- ALSHAWI H. (1987). *Memory and Context for Language Interpretation*. Cambridge University Press, Cambridge
- BEUN R.J., CREMERS A.H.M. (1998). « Object Reference in a Shared Domain of Conversation ». In *Pragmatics and Cognition* 6 (1/2).
- CARON J. (1989). *Précis de psycholinguistique*. PUF, Paris.
- CORBLIN F. (1987). *Indéfini, défini et démonstratif*. Droz, Genève.
- GROBET A. (2002). *L'identification des topiques dans les dialogues*. De Boeck-Duculot, Bruxelles.
- GROSZ B., JOSHI A., WEINSTEIN S. (1995). « Centering: A Framework for Modeling the Local Coherence of Discourse ». In *Computational Linguistics* 21 (2) : 203-225.
- HAJIČOVÁ E., HOSKOVEC T., SGALL P. (1995). « Discourse Modelling Based on Hierarchy of Saliency ». In *Prague Bulletin of Mathematical Linguistics* 64 : 5-24.
- ISARD S. (1975). « Changing the Context ». In E.L. KEENAN (éd.), *Formal Semantics of Natural Language*. Cambridge University Press, London & New York : 287-296.
- KAMP H., REYLE U. (1993). *From Discourse to Logic*. Kluwer, Dordrecht.
- LANDRAGIN F. (2004). « Saillance physique et saillance cognitive ». In *Cognition, Représentation, Langage (CORELA)* 2 (2), <http://edel.univ-poitiers.fr/corela/>.
- LANDRAGIN F. (2005). « Traitement automatique de la saillance ». In *Actes de TALN 2005*. Dourdan : 263-272.
- LAPPIN S., LEASS H. (1994). « A Syntactically Based Algorithm for Pronominal Anaphora Resolution ». In *Computational Linguistics* 20 (4) : 535-561.
- MITKOV R. (1998). « Robust Pronoun Resolution with Limited Knowledge ». In *Proceedings of the 18th International Conference on Computational Linguistics*. Montréal.

- SIDNER C. (1979). *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*. Ph.D. thesis, Massachusetts Institute of Technology.
- WEIL-BARAIS A. (éd.) (1993). *L'homme cognitif*. PUF, Paris.